

Extinction of metastable stochastic populations

Michael Assaf and Baruch Meerson

Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem 91904, Israel

We investigate the phenomenon of extinction of a long-lived self-regulating stochastic population, caused by intrinsic (demographic) noise. Extinction typically occurs via one of two scenarios depending on whether the absorbing state $n = 0$ is a repelling (scenario A) or attracting (scenario B) point of the *deterministic rate equation*. In scenario A the metastable stochastic population resides in the vicinity of an attracting fixed point next to the repelling point $n = 0$. In scenario B there is an intermediate repelling point $n = n_1$ between the attracting point $n = 0$ and another attracting point $n = n_2$ in the vicinity of which the metastable population resides. The crux of the theory is a dissipative variant of WKB (Wentzel-Kramers-Brillouin) approximation which assumes that the typical population size in the metastable state is large. Starting from the master equation, we calculate the quasi-stationary probability distribution of the population sizes and the (exponentially long) mean time to extinction for each of the two scenarios. When necessary, the WKB approximation is complemented (i) by a recursive solution of the quasi-stationary master equation at small n and (ii) by the van Kampen system-size expansion, valid near the fixed points of the deterministic rate equation. The theory yields both entropic barriers to extinction and pre-exponential factors, and holds for a general set of multi-step processes when detailed balance is broken. The results simplify considerably for single-step processes and near the characteristic bifurcations of scenarios A and B.

PACS numbers: 02.50.Ga, 87.23.Cc

I. INTRODUCTION

Extinction of an isolated stochastic population after maintaining a long-lived state is a dramatic phenomenon. It occurs, even in the absence of environmental variations, because of an unusual chain of random events when population losses dominate over gains. Population extinction risk is a key negative factor in viability of small populations [1, 2], whereas extinction of a disease following an epidemic outburst [1, 3] is of course favorable. The possibility and consequences of extinction of biologically important components, regulated by chemical reactions in living cells, have also attracted interest [4]. As stochastic population dynamics are usually far from equilibrium, and no general methods of evaluating large fluctuations are available, they are of much interest to physics [5, 6].

This work deals with an isolated single-species population undergoing a set of gain-loss processes. We will assume that the population is well mixed, so that spatial degrees of freedom are irrelevant. At the level of the *deterministic rate equation* (henceforth *rate equation*), which describes the time history of the mean population size $\bar{n}(t)$ and ignores fluctuations, $\bar{n}(t)$ flows to an attracting fixed point, where the gain and loss processes balance each other. The actual stochastic population, however, behaves differently and ultimately becomes extinct. This is because, in the absence of influx of new individuals, the empty state $n = 0$ is *absorbing*: the probability of exiting from it is zero [7].

Although extinction (and fluctuations in general) are beyond its scope, the rate equation is a convenient starting point of our analysis. For an isolated single-species population the rate equation can be written as

$$\frac{d\bar{n}}{dt} = \bar{n} \Phi(\bar{n}), \quad (1)$$

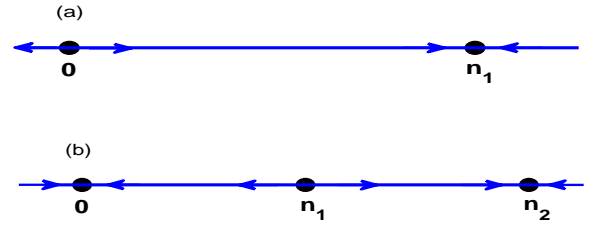


FIG. 1: (color online). Typical extinction scenarios are determined by the character of the fixed point $\bar{n} = 0$ of the rate equation (1). (a) Scenario A: the fixed point $\bar{n} = 0$ is repelling. In the *stochastic* system extinction occurs via a large fluctuation which brings the metastable population from a vicinity of the attracting fixed point $\bar{n} = n_1$ of the rate equation directly to the absorbing state $n = 0$. (b) Scenario B: the fixed point $\bar{n} = 0$ is attracting. In the stochastic system extinction occurs via a large fluctuation which brings the metastable population from a vicinity of the *next* attracting fixed point $\bar{n} = n_2$ of the rate equation to a vicinity of the repelling fixed point $\bar{n} = n_1$. From there the population flows “downhill” to the absorbing state $n = 0$ almost deterministically.

where $\Phi(x)$ is a smooth function determined by the specific gain-loss processes, see below. For generic gain-loss processes $\Phi'(0) \neq 0$. For $\Phi'(0) > 0$ the fixed point $\bar{n} = 0$ is repelling, whereas for $\Phi'(0) < 0$ it is attracting. In the former case, the next fixed point $\bar{n} = n_1 > 0$ of Eq. (1) is attracting, see Fig. 1a. According to the rate equation, the mean population size in this case flows to $\bar{n} = n_1$ and stays there forever. When varying the rate constants of the gain-loss processes, the attracting fixed point $\bar{n} = n_1$ emerges via a transcritical bifurcation.

Now let $\bar{n} = 0$ be an attracting fixed point of the rate equation (1). To have a long-lived population of

a nonzero size, at least two more fixed points of the rate equation (1) must be present: a repelling point $\bar{n} = n_1 > 0$ and an attracting point $\bar{n} = n_2 > n_1$, see Fig. 1b. When starting from any $\bar{n}(t=0) > n_1$, the mean population size flows to $\bar{n} = n_2$ and, according to the rate equation, stays there forever. The characteristic bifurcation in this case is saddle-node.

As we will see shortly, these two cases give rise to two different extinction scenarios of stochastic populations. To account for the intrinsic noise, we employ the master equation

$$\frac{dP_n(t)}{dt} = \sum_r [W_r(n-r)P_{n-r}(t) - W_r(n)P_n(t)] \quad (2)$$

which describes the evolution of the probability $P_n(t)$ to have n individuals at time t . Here $W_r(k) \geq 0$ is the transition rate between the states with k and $k+r$ individuals, whereas $r = \pm 1, \pm 2, \dots$, and all terms that include P_k with $k < 0$ are assumed to be zero. For $P_0(t)$ the master equation is

$$\frac{dP_0(t)}{dt} = \sum_{r < 0} W_r(-r)P_{-r}(t). \quad (3)$$

For $n = 0$ to be an absorbing state, the process rates must obey, for any $r = \pm 1, \pm 2, \dots$, the condition $W_r(0) = 0$.

We will be interested in the important regime of parameters for which the mean population size in the metastable state, as predicted by Eq. (1), is large compared to one. Here, prior to extinction, a long-lived probability distribution function (PDF) of the population sets in, on a relaxation time scale t_r , around the corresponding attracting fixed point of the rate equation. This long-lived PDF, however, is metastable: it slowly decays in time. Simultaneously, the probability to find the population extinct slowly grows in time, see e.g. Refs. [8, 9]:

$$P_{n>0}(t \gg t_r) \simeq \pi_n e^{-t/\tau}, \quad P_0(t \gg t_r) \simeq 1 - e^{-t/\tau}. \quad (4)$$

The shape function π_n ($n = 1, 2, \dots$) of the metastable PDF is called the quasi-stationary distribution (QSD). For metastable populations a very strong inequality, $\tau \gg t_r$ holds, and the decay time τ is equal to the mean time to extinction (MTE): the mean time it takes the stochastic process to reach the absorbing state at $n = 0$. The main objectives of this work is to accurately, and analytically, calculate the QSD π_n and the MTE τ of a population which experiences quite a general set of stochastic gain-loss processes. The crux of the method is a dissipative WKB approximation [10–12], where one assumes $n \gg 1$, treats n as a continuous variable and searches for π_n as

$$\pi_n = e^{-NS(n) - S_1(n) - (1/N)S_2(n) - \dots}. \quad (5)$$

Here $N \gg 1$ is a large parameter which scales as the mean population size in the metastable state. $S(n)$ is called the action, whereas $a(n) = e^{-S_1(n)}$ is called the

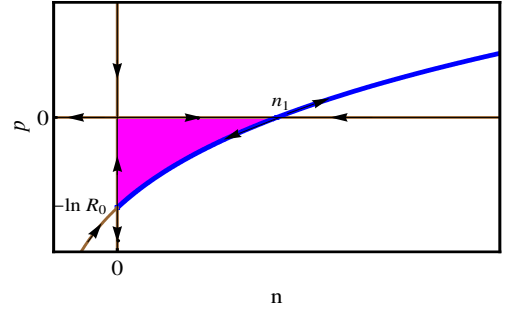


FIG. 2: (color online). Example of scenario A of population extinction driven by intrinsic noise. Shown are the zero-energy trajectories of the WKB Hamiltonian $H(n, p)$ for the reactions $A \xrightarrow{\lambda} 2A$, $A \xrightarrow{\mu} \emptyset$ and $2A \xrightarrow{\sigma} \emptyset$ [16]. The trajectories denoted by the thicker line determine the WKB solution for π_n , obtained in Ref. [16]. The activation trajectory connects the metastable point $(n_1, 0)$ and the fluctuational extinction point $(0, p_f)$, where $p_f = -\ln R_0$, and $R_0 = \lambda/\mu$. The effective entropy barrier to extinction is equal to $N\Delta S$, where $N = \lambda/\sigma$ and ΔS is the area of the shaded region, given by Eq. (76).

amplitude. The WKB approximation breaks down at $n = \mathcal{O}(1)$. Here a different approximation must be used, as explained below.

Here is an overview of the two extinction scenarios as described by the WKB approximation. First, let $\bar{n} = 0$ be a repelling fixed point of the rate equation, see Fig. 1a. In a stochastic description extinction occurs via a large fluctuation which, acting against an effective entropy barrier, brings the population from a vicinity of $n = n_1$ directly to the absorbing state $n = 0$. In the WKB language this transition is possible because of the presence of the fluctuational momentum $p = dS/dn$, see Fig. 2. The attracting and repelling fixed points of the rate equation $\bar{n} = n_1$ and $\bar{n} = 0$, respectively, become hyperbolic fixed points of an extended phase plane (n, p) . Importantly, an additional hyperbolic fixed point $(0, p_f)$ - the fluctuational extinction point - appears here, with a zero coordinate, $n = 0$, but a nonzero momentum p_f [13–15]. The most probable path to extinction is the heteroclinic trajectory, directly connecting the “metastable point”, that is the hyperbolic point $(n_1, 0)$, and the “fluctuational extinction point”: the hyperbolic point $(0, p_f)$. (Such escape trajectories - heteroclinic trajectories with a non-zero momentum - are often called “activation trajectories”, see e.g. [12].) This is what we call extinction scenario A.

Now let $\bar{n} = 0$ be an attracting fixed point of the rate equation (1), so that the metastable population resides in the vicinity of $n = n_2$, see Fig. 1b. In the stochastic description, extinction occurs via a large fluctuation which brings the population from a vicinity of $n = n_2$ to a vicinity of the *repelling* fixed point n_1 . From there the system flows into the absorbing state $n = 0$ “downhill”, that is almost deterministically. In the framework of WKB theory the transition from n_2 to n_1 occurs in the extended

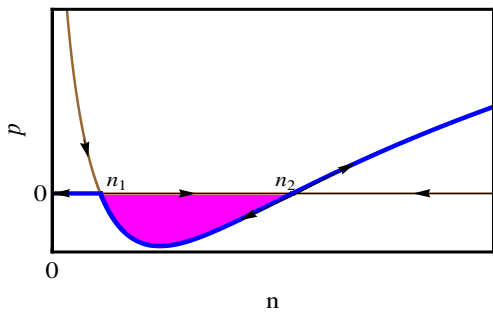


FIG. 3: (color online). Example of scenario B of population extinction driven by intrinsic noise. Shown are the zero-energy trajectories (106) on the phase plane (n, p) of the WKB Hamiltonian (105) for the reactions $A \xrightarrow{\mu} \emptyset$, $2A \xrightarrow{\lambda} 3A$ and $3A \xrightarrow{\sigma} 2A$. The trajectories denoted by the thicker lines determine the WKB solution for the QSD. The most probable path to extinction first goes along the non-zero-momentum heteroclinic trajectory (the activation trajectory) which connects the points $(n_2, 0)$ and $(n_1, 0)$. Then the population flows almost deterministically to the extinction point $(0, 0)$ along a zero-momentum segment (the relaxation trajectory). The effective entropy barrier to extinction is equal to $N\Delta S$, where $N = 3\lambda/(2\sigma)$, and ΔS is the area of the shaded region, given by Eq. (110).

phase plane (n, p) where all three fixed points are hyperbolic, see Fig. 3. Here the optimal path to extinction is composed of two segments: the non-zero-momentum heteroclinic trajectory connecting the hyperbolic fixed points $(n_2, 0)$ and $(n_1, 0)$ (the activation trajectory), and the zero-momentum segment going from $n = n_1$ to $n = 0$ (the relaxation trajectory). This is what we call extinction scenario B.

The mean time to extinction (MTE) τ and/or the QSD π_n of metastable single-species stochastic populations were calculated previously in particular examples in different contexts of physics, chemistry, population biology, epidemiology, cell biology, *etc.* Among them there is a large body of work which approximated the master equation by an effective Fokker-Planck equation, derived via the van Kampen system size expansion or related recipes. Once the Fokker-Planck equation is obtained, the MTE and QSD can be calculated by standard methods [5, 6]. Unfortunately, this approximation is in general uncontrolled. It fails in its description of the tails of the QSD, and gives exponentially large errors in the MTE, as shown in Refs. [16–19].

With a few exceptions, accurate analytic results for the MTE and QSD are only available for *single-step* gain-loss processes: $r = \pm 1$ in Eq. (2). In this case the MTE can be determined exactly by employing the backward master equation [5, 6]. This yields a cumbersome analytic expression for the MTE which, for a large population size in the metastable state, can be simplified via a saddle-point approximation. Such a procedure was implemented in Ref. [18]. In its turn, the QSD π_n of single-step processes can be calculated from a recursive

relation obtained when substituting Eq. (4) in the master equation. Several model examples of *multi-step* processes were considered in Refs. [9, 16, 19, 20], all of them belonging to extinction scenario A. We will generalize the previous results substantially and determine the MTE and QSD for quite a general set of gain-loss processes pertaining to extinction scenario A. We will also determine the MTE and QSD for extinction scenario B.

Our WKB theory starts with applying the ansatz (5) to an eigenvalue problem for the QSD π_n which is nothing but the first excited eigenvector of the master equation. In the leading WKB order one arrives at the problem of finding zero-energy trajectories of an effective classical Hamiltonian [12]. There are two different types of zero-energy phase trajectories (in addition to the extinction line $q = 0$): the activation and relaxation trajectories, which correspond to the fast and slow WKB modes, respectively [21]. To obtain the pre-exponents, one needs to consider the sub-leading WKB order. The WKB calculations are simpler for scenario A, as the relaxation trajectory does not play any role here. In scenario B both the activation, and the relaxation trajectories are important. In both scenarios the WKB approximation breaks down at $n = \mathcal{O}(1)$. Here we find the QSD, up to a normalization constant, from a recursive relation, obtained by linearizing the process rates with respect to n at sufficiently small n . In scenario A it suffices to match the recursive solution with the fast-mode solution in their joint region of validity, in much the same way as it was done by Kessler and Shnerb [16] in a particular example of three stochastic reactions. In scenario B the slow mode dominates the WKB-solution at $n < n_1$. It diverges, however, at $n = n_1$. To obtain a regular solution there, one needs to go beyond the WKB approximation and account, in a close vicinity of $n = n_1$, for strong coupling between the fast- and slow-mode solutions. This can be done via the van-Kampen system size expansion of the master equation which does hold in the vicinity of $n = n_1$. This procedure was first implemented by Meerson and Sasorov [21], in a model problem of noise-driven population explosion. Then it has been employed by Escudero and Kamenev [22] in the context of a WKB theory of stochastic population switches between two different metastable states. The theory of Escudero and Kamenev [22] was formulated for quite a general set of gain-loss processes. In this paper we will adopt their general approach, and some of their notation, in the problem of population extinction.

Here is a plan of the remainder of the paper. Section II starts with a formulation of the eigenvalue problem for the QSD. Then we expose the WKB approximation and the fast- and slow-mode WKB solutions. The derivation here is quite general and holds for extinction scenarios A and B. Section III presents a derivation of recursive solution of the quasi-stationary master equation for sufficiently small n . This derivation, which also holds for extinction scenarios A and B, is specific to population extinction. Except for simple particular cases (see *e.g.* Ref.

[16]), it has not been attempted before. In Section IV we match the fast-mode WKB solution with the recursive small- n solution and obtain general expressions for the QSD and MTE in scenario A. In the same Section we obtain the QSD and MTE for single-step processes and near the transcritical bifurcation, characteristic of scenario A. Then we illustrate our theory on several particular examples, some of which investigated previously. In Section V we determine the QSD and MTE for scenario B. Then we again apply our results to single-step processes and near the saddle-node bifurcation, characteristic of scenario B. Furthermore, we consider a particular example of three stochastic reactions and compare our theoretical predictions with a numerical solution of the master equation. A summary of our results is presented in Section VI.

II. EIGENVALUE PROBLEM, WKB APPROXIMATION, AND FAST- AND SLOW-MODE SOLUTIONS

When starting at $t = 0$ from a sufficiently large population, the probability distribution $P_n(t)$, as described by the master equation (2) approaches, on a relaxation time scale t_r , a long-lived metastable PDF peaked at a non-zero attracting fixed point of the rate equation. The metastable distribution is slowly “leaking” to zero, see Eq. (4). Let us denote the non-zero attracting fixed point by $n = n_*$ ($n_* = n_{1,2}$ for scenario A and B, respectively, see Figs. 2 and 3). Using Eq. (4), we arrive at an eigenvalue problem for the QSD π_n , $n = 1, 2, \dots$:

$$\sum_r [W_r(n-r)\pi_{n-r} - W_r(n)\pi_n] = -E\pi_n. \quad (6)$$

Importantly, the eigenvalue $E = 1/\tau$ turns out to be *exponentially* small compared to the relaxation time t_r . Therefore, the term in the right hand side of Eq. (6) can be neglected [12, 16, 21, 22], and we have to deal with a quasi-stationary equation

$$\sum_r [W_r(n-r)\pi_{n-r} - W_r(n)\pi_n] = 0, \quad n = 1, 2, \dots \quad (7)$$

For definiteness, we normalize the QSD to unity: $\sum_{n=1}^{\infty} \pi_n = 1$. Once π_n is found, we can use Eqs. (3) and (4) to calculate the MTE:

$$E = 1/\tau = \sum_{r < 0} W_r(-r)\pi_{-r}. \quad (8)$$

Let us introduce a rescaled coordinate $q = n/N$, where $N \gg 1$ is the large parameter of the problem. The central assumption of our theory is that, after a proper rescaling of time which will be introduced shortly, the process rates can be represented as

$$W_r(n) \equiv W_r(Nq) = Nw_r(q) + u_r(q) + \mathcal{O}(1/N), \quad (9)$$

where, for $q = \mathcal{O}(1)$, $w_r(q)$ and $u_r(q)$ are $\mathcal{O}(1)$. This assumption guarantees that the population be long-lived,

and is crucial both for the WKB-approximation that we present in this Section, and for the recursive solution of Eq. (7) that we will be dealing with later. As $q = 0$ is the absorbing state, $w_r(0) = u_r(0) = 0$.

For $n \gg 1$ we can employ the WKB ansatz (5):

$$\pi(q) \equiv \pi_{Nq} \simeq Ae^{-NS(q) - S_1(q)}, \quad (10)$$

where $S(q)$ and $S_1(q)$ are assumed to be $\mathcal{O}(1)$, and a constant prefactor A is introduced for convenience, see below. Now we assume that $|r_{max}| \ll N$, Taylor-expand the functions of $q - r/N$ in Eq. (7) around q and keep terms up to $\mathcal{O}(1)$ order. We obtain the equation derived by Escudero and Kamenev [22]:

$$\sum_r (Nw_r + u_r) \times \left[e^{rS'} \left(1 + \frac{r}{N} S'_1 - \frac{r^2}{2N} S'' - \frac{r}{N} \frac{w'_r}{w_r} \right) - 1 \right] = 0, \quad (11)$$

where the primes denote differentiation with respect to q . In the leading order $\mathcal{O}(N)$, this equation yields a stationary Hamilton-Jacobi equation $H(q, S') = 0$, where

$$H(q, p) = \sum_r w_r(q) (e^{rp} - 1) \quad (12)$$

is the effective Hamiltonian, and $p = S'$ is the momentum [12]. Therefore, in the leading WKB order, one needs to find *zero-energy* phase trajectories of the Hamiltonian (12). As $w_r(0) = 0$ for any r , one such trajectory is $q = 0$ at an arbitrary p : the extinction line. This line is of no importance in the WKB theory, however. What we need are phase trajectories $p = p(q)$. One of them is the relaxation trajectory $p = 0$. In general, there is one and only one additional phase trajectory for which $p = p_a(q) \neq 0$, except in some points q . Let us prove this statement. The Hamiltonian $H(q, p)$ vanishes at $p = 0$. Differentiating Eq. (12) twice with respect to p , we obtain $H_{pp}(q, p) = \sum_r r^2 w_r(q) e^{rp} > 0$. Therefore H is a convex function of p , and so it has one and only one additional real root [23]. The relaxation trajectory $p = 0$ and activation trajectory $p = p_a(q)$ give rise to the slow and fast WKB modes, respectively, as was shown in a particular example in Ref. [21].

The q -dynamics along $p = 0$ is described by the Hamilton's equation

$$\dot{q} = \frac{\partial H(q, p)}{\partial p} \Big|_{p=0} \equiv H_p(q, 0) = \sum_r r w_r(q) \quad (13)$$

which is nothing but the (rescaled) rate equation (1). The nontrivial fixed points of the rate equation, $q_i = n_i/N$, are positive roots of the equation $H_p(q_i, 0) = 0$. As $p_a(q_i) = 0$, the activation trajectory $p = p_a(q)$ crosses the relaxation trajectory in these fixed points. As $w_r(0) = 0$ for any r , the small- q expansion of Eq. (13) generically starts with a linear term in q . In the remainder of this paper we assume that the linear decay rate is non-zero in

the leading order: $w'_{-1}(0) = \alpha > 0$. Therefore, we can always rescale time, and all the rates, by α . This procedure uniquely defines the rescaling leading to Eq. (9).

In extinction scenario A, see Fig. 2, the most probable path to extinction is the heteroclinic trajectory connecting the fixed point $(q_*, 0)$ of the Hamiltonian system (here q_* is the attracting point of the rate equation) with the fluctuational extinction point $(0, p_f)$.

In extinction scenario B, see Fig. 3, the most probable path to extinction is composed of two segments. The first one is the activation trajectory: a non-zero-momentum heteroclinic trajectory connecting the fixed point $(q_*, 0)$ with an intermediate fixed point $(q_{rep}, 0)$, where q_{rep} is a repelling fixed point of the rate equation. The second one is a relaxation segment $p = 0$, connecting the point $(q_{rep}, 0)$ with the deterministic extinction point $(0, 0)$.

For the fast mode one obtains [12]

$$S(q) = S^{(f)}(q) = \int^q d\xi p_a(\xi), \quad (14)$$

where the integration constant is already accounted for by the prefactor A in Eq. (10). For the slow mode $S(q) = S^{(s)}(q) = 0$.

In the subleading $\mathcal{O}(1)$ order, Eq. (11) yields a first-order differential equation for $S_1(q)$:

$$\sum_r w_r e^{rp} \left(r S'_1 - \frac{r^2}{2} p' - \frac{r w'_r}{w_r} \right) + u_r (e^{rp} - 1) = 0, \quad (15)$$

where $p = p_a(q)$ for the fast mode, and $p = 0$ for the slow mode. It is convenient to use the identities

$$\begin{aligned} H_q &= \sum_r w'_r(q) (e^{rp} - 1), & H_p &= \sum_r r w_r(q) e^{rp}, \\ H_{qq} &= \sum_r w''_r(q) (e^{rp} - 1), & H_{pp} &= \sum_r r^2 w_r(q) e^{rp}, \\ H_{qp} &= \sum_r r w'_r(q) e^{rp}, \end{aligned} \quad (16)$$

where the subscripts p and q stand for the partial derivatives. To remind the reader, all the rates in Eqs. (16) are rescaled with respect to the linear decay rate constant α . The fast-mode solution for $S_1(q)$ can be written as

$$\begin{aligned} S_1^{(f)}(q) &= \int^q d\xi \\ &\times \frac{H_{qp}(\xi, p_a) + \frac{1}{2} H_{pp}(\xi, p_a) p'_a(\xi) - \sum_r u_r(\xi) (e^{rp_a} - 1)}{H_p(\xi, p_a)}, \end{aligned} \quad (17)$$

where $p_a = p_a(\xi)$. This result was obtained by Escudero and Kamenev [22] in the context of stochastic population switches. The quantity $H_p(\xi, p_a)$ in the denominator of the integrand in Eq. (17) vanishes in every fixed point $\xi = q_i$ of the rate equation, including $\xi = 0$. To see how the integrand behaves at the fixed points, consider

the equation $H[q, p_a(q)] = 0$ for the activation trajectory. Differentiating it with respect to q , we obtain

$$H_q[q, p_a(q)] + H_p[q, p_a(q)] p'_a(q) = 0. \quad (18)$$

One more differentiation gives

$$H_{qq} + H_{pp} p''_a + (2H_{pq} + H_{pp} p'_a) p'_a = 0, \quad (19)$$

evaluated at $p = p_a(q)$. By virtue of identities (16) each of the first two terms of Eq. (19) vanishes at $q = q_i$ and $p = p_a(q_i) = 0$, so the expression $2H_{pq} + H_{pp} p'_a$ must also vanish there. As a result, the first two terms in the numerator of the integrand of Eq. (17) cancel each other at $\xi = q_i$. The remaining term in the numerator, $-\sum_r u_r(\xi) (e^{rp_a(\xi)} - 1)$, is proportional to $(\xi - q_i)$ in the vicinity of $\xi = q_i$, exactly as the quantity $H_p(\xi, p_a)$ in the denominator. Therefore, the integrand is well behaved at $\xi = q_i > 0$. At $q = 0$, $S_1^{(f)}(q)$ diverges logarithmically, see below. This divergence does not cause any concern, as the WKB approximation (which demands $n \gg 1$, or $q \gg N^{-1}$, for its validity) does not hold for small q anyway.

One can partially perform the integration over q in Eq. (17) by using Eqs. (18) and (19). After some algebra,

$$S_1^{(f)}(q) = -\ln \sqrt{|S''(q)|} + \Psi(q), \quad (20)$$

where

$$\Psi(q) = \int^q \left[\frac{H_{qq}(\xi, p_a)}{2H_q(\xi, p_a)} - \frac{\sum_r u_r(\xi) (e^{rp_a} - 1)}{H_p(\xi, p_a)} \right] d\xi, \quad (21)$$

and $S''(q) \equiv p'_a(q)$. Note that $S''(q)$ does not vanish in any of the nontrivial fixed points q_i , so the logarithmic term in Eq. (20) is well behaved there [24]. Indeed, from Eq. (18) $S''(q) = p'_a(q) = -H_q[q, p_a(q)]/H_p[q, p_a(q)]$. By Taylor-expanding the numerator and denominator in the vicinity of any nontrivial fixed point q_i , one can see that $S''(q_i) \neq 0$.

Therefore, the general fast-mode solution is given by Eq. (10) with $S^{(f)}(q)$ from Eq. (14), and $S_1^{(f)}(q)$ from Eqs. (20) and (21). It also includes a constant prefactor $A = A_f$ which can be found immediately. Indeed, at $N \gg 1$ the QSD is strongly peaked around the attracting fixed point $q = q_*$. Here the fast-mode solution dominates, and A_f can be found by normalizing to unity the gaussian asymptote of the QSD around $q = q_*$. The gaussian asymptote is obtained by expanding the QSD (10) in the vicinity of $q = q_*$:

$$\pi(q) \simeq A_f e^{-N S^{(f)}(q_*) - S_1^{(f)}(q_*) - (N/2) S''(q_*) (q - q_*)^2}, \quad (22)$$

where we have used the equalities $S'(q_*) = p_a(q_*) = 0$. To see that $S''(q_*) > 0$, one can again use Eq. (19). At $q = q_*$ it reads $2H_{pq}(q_*, 0) + H_{pp}(q_*, 0) p'_a(q_*) = 0$. By virtue of Eq. (16) $H_{pp}(q_*, 0) > 0$. The quantity $H_{pq}(q_*, 0)$ is the q -derivative, evaluated at $q = q_*$, of the expression in the right hand side of the rate equation (13). As the point $q = q_*$ is by assumption attracting, $H_{pq}(q_*, 0) < 0$.

Therefore, $p'_a(q_*) = S''(q_*) > 0$, and the asymptote (22) is indeed a gaussian distribution. Normalizing it to unity, we obtain

$$A_f = \sqrt{\frac{S''(q_*)}{2\pi N}} e^{N S^{(f)}(q_*) + S_1^{(f)}(q_*)}, \quad (23)$$

so the fast-mode solution is fully determined:

$$\pi(q) = \sqrt{\frac{S''(q_*)}{2\pi N}} e^{N[S^{(f)}(q_*) - S^{(f)}(q)] + S_1^{(f)}(q_*) - S_1^{(f)}(q)}, \quad (24)$$

with $S^{(f)}(q)$ from Eq. (14) and $S_1^{(f)}(q)$ from Eqs. (20) and (21).

Now consider the slow-mode solution for which $S^{(s)}(q) = 0$. The subleading-order contribution $S_1^{(s)}(q)$ is found by putting $p = 0$ in Eq. (17):

$$S_1^{(s)}(q) = \int^q d\xi \frac{H_{pq}(\xi, 0)}{H_p(\xi, 0)} = \ln H_p(q, 0). \quad (25)$$

Then Eq. (10) yields the general slow-mode solution for the QSD [21, 22]:

$$\pi_s(q) = -\frac{A_s}{H_p(q, 0)}, \quad (26)$$

where A_s is an arbitrary constant. The minus sign is put here for convenience, because in the region of $0 < q < q_1$, where the slow-mode solution is relevant (see Section V), $\dot{q} = H_p(q, 0) < 0$ and $A_s > 0$. One can see from Eq. (26) that the slow-mode solution diverges in the fixed points of the rate equation [21]. This divergence will be cured in Section V.

We show in the following that, for a given extinction scenario and in a given region of q , only one of the modes, either fast or slow, dominates the resulting QSD, while the other one must be discarded. Before we deal with this issue, however, we recall that the WKB approximation breaks down at $n = \mathcal{O}(1)$. To find the QSD for *all* n we will solve Eq. (7) in the region of $n \ll N$ by recursion and then match the recursive solution with either the fast-mode (in scenario A), or the slow-mode (in scenario B) WKB solution in the joint region of their validity.

III. RECURSIVE SOLUTION

The objective of this Section is to approximately solve Eq. (7) at sufficiently small n . The exact criterion of smallness will appear later, when we match different solutions in joint regions of their validity.

In the leading order in N we take $W_r(n) = N w_r(n/N)$, see Eq. (9), and expand it in n/N up to the linear term: $W_r(n) \simeq N w_r(0) + n w'_r(0) = n w'_r(0)$. Then Eq. (7) becomes

$$\sum_r w'_r(0) [(n-r)\pi_{n-r} - n\pi_n] = 0, \quad (27)$$

where only processes with $w'_r(0) \neq 0$ contribute. One can look for particular solutions of this recursive equation in the form $\pi_n = f_n/n$ thus arriving at an equation with n -independent coefficients:

$$\sum_r w'_r(0) (f_{n-r} - f_n) = 0. \quad (28)$$

In the remainder of this paper we make the following simplifying assumption:

$$w'_{-2}(0) = w'_{-3}(0) = \dots = 0. \quad (29)$$

That is, we assume that the rates of the multi-step *loss* processes $n \rightarrow n-m$, where $m = 2, 3, \dots$, do not have, in the leading order in N , linear terms in their Taylor expansion in n . This assumption is always satisfied for stochastic chemical reactions (where pairs, triplets, ..., of reacting particles are needed to bring down the number of particles by 2, 3, ...). The conditions (29) also hold for all models of population biology and epidemiology we are aware of.

Using Eq. (29) and the equality $w'_{-1}(0) = 1$ (to remind the reader, we are using rescaled variables), we rewrite the recursive Eq. (28) as

$$f_{n+1} = \left[1 + \sum_{r=1}^K w'_r(0) \right] f_n - \sum_{r=1}^K w'_r(0) f_{n-r}, \quad (30)$$

where $K \equiv r_{max} \ll N$. If there is no degeneracy, the general solution of Eq. (30) is a linear combination of all particular solutions $f_n = \lambda^{-n}$, where λ obeys the characteristic polynomial equation of degree $K+1$:

$$\sum_{r=1}^K w'_r(0) \lambda^{r+1} - \left[1 + \sum_{r=1}^K w'_r(0) \right] \lambda + 1 = 0. \quad (31)$$

Note, that $\lambda = \lambda_0 = 1$ is always a root. Let us show that Eq. (31) has one and only one additional positive root, λ_1 , while all others roots $\lambda_2, \lambda_3, \dots, \lambda_K$ are either negative or complex. First, we establish a connection between the roots λ_i of Eq. (31) and the crossing points with the p -axis of the zero-energy trajectories of the WKB Hamiltonian (12). By expanding $w_r(q \rightarrow 0) \simeq w'_r(0)q$, Eq. (12) becomes

$$\sum_r w'_r(0) (e^{rp} - 1) = 0, \quad (32)$$

where only terms with $w'_r(0) \neq 0$ contribute. Putting $e^p = \lambda$ and $w'_{-1}(0) = 1$ and using Eq. (29), one can see that Eq. (32) coincides with Eq. (31). As we have shown that the equation $H(q, p) = 0$ has, for any q , two and only two real solutions for p [23], Eq. (31) also has two and only two real solutions for λ , both of them positive. The roots $\lambda_0 = 1$ and λ_1 correspond to the crossing points with the p -axis of the slow and fast modes, respectively.

Dividing Eq. (31) by $1 - \lambda$, we arrive at a polynomial equation of degree K :

$$1 - \sum_{r=1}^K w'_r(0) (\lambda + \dots + \lambda^r) = 0. \quad (33)$$

For $K \leq 4$ the roots of this polynomial can be expressed in radicals. For $K \geq 5$, they need to be computed numerically. Assume that we have found all of the roots of Eq. (33) λ_i , $i = 1, 2, \dots, K$. If there is no degeneracy, the general solution for f_n is

$$f_n = \sum_{i=0}^K C_i \lambda_i^{-n} = C_0 + \sum_{i=1}^K C_i \lambda_i^{-n}, \quad (34)$$

where the coefficients C_i are the following (see Appendix A for the derivation):

$$C_i = \frac{(-1)^K f_1 \prod_{j=1}^K \lambda_j}{\prod_{\substack{j=0 \\ j \neq i}}^K (\lambda_i - \lambda_j)}. \quad (35)$$

The coefficient C_0 , corresponding to the root $\lambda_0 = 1$, can be expressed through the coefficients $w'_r(0)$, $r = 1, 2, \dots, K$ (see Appendix A):

$$C_0 = \frac{f_1}{1 - w'_1(0) - 2w'_2(0) - \dots - Kw'_K(0)}. \quad (36)$$

Before writing down the general solution of the recursive equation (27) for the QSD, we recall a simple relation [25] between f_1 and the MTE τ . Using the rescaled reaction rates, we can rewrite Eq. (8) as

$$\tau^{-1} = \alpha \sum_{r < 0} W_r(-r) \pi_{-r}.$$

In view of the conditions (29), only one term in the sum survives in the leading order. As $w'_{-1}(0) = 1$, we obtain

$$\tau^{-1} = \alpha W_{-1}(1) \pi_1 \simeq \alpha \pi_1 = \alpha f_1,$$

so $f_1 \simeq 1/(\alpha\tau)$. Now we switch to the rescaled variable $q = n/N$, use the relation $\pi(q) = f_n/(Nq)$ and Eq. (34), and obtain the small- q asymptote of $\pi(q)$:

$$\pi(q) = \frac{(-1)^K \prod_{j=1}^K \lambda_j}{\alpha\tau Nq} \sum_{i=0}^K \frac{\lambda_i^{-Nq}}{\prod_{\substack{j=0 \\ j \neq i}}^K (\lambda_i - \lambda_j)}. \quad (37)$$

The validity region of this asymptote (which includes the yet unknown τ to be found later) is scenario-dependent. It is a relatively narrow region $q \ll N^{-1/2}$ in scenario A, and a broader region $q \ll 1$ in scenario B. The difference comes from the fact that in scenario A the recursive solution needs to be matched, at $n \gg 1$, with a rapidly growing fast-mode solution, whereas in scenario B the

matching needs to be done with a slowly varying slow-mode solution, see Sections IV and V, respectively.

What is the role of complex roots of the polynomial equation (33) in the recursive solution (37)? These can appear only for $K \geq 3$, and they come in complex conjugate pairs: λ_j and $\lambda_k = \overline{\lambda_j}$. One can show, by using Eq. (35), that the coefficients C_j and C_k , corresponding to λ_j and $\overline{\lambda_j}$, are also complex conjugate: $C_k = \overline{C_j}$, so that $\pi(q)$ from Eq. (37) is real-valued as expected. When complex roots are present, the QSD at small n may exhibit rapidly decaying oscillations as a function of n .

Let us now determine the $n \gg 1$, or $q \gg N^{-1}$, asymptote of the QSD (37), in each of the two extinction scenarios. This asymptote will be matched, in each scenario, with the dominant WKB mode. Equation (37) includes $K + 1$ terms. At $n \gg 1$ the leading contribution comes from the term with the smallest $|\lambda_i|$. The rest of the terms are exponentially small compared to the leading one and can be safely neglected.

In scenario A, the two positive roots of Eq. (31) are $0 < \lambda_1 < 1$ and $\lambda_0 = 1$, whereas the rest of the (negative or complex) roots obey the inequality $|\lambda_{i>1}| > \lambda_1$, see Appendix B. In this case the asymptote of the recursive solution (37) at $n \gg 1$, or $q \gg N^{-1}$, is

$$\pi(q) \simeq \frac{A_1 \lambda_1^{-Nq}}{\alpha\tau Nq}, \quad (38)$$

where the positive constant A_1 (see Appendix B) satisfies

$$A_1 = \frac{(-1)^K \prod_{j=1}^K \lambda_j}{\prod_{\substack{j=0 \\ j \neq 1}}^K (\lambda_1 - \lambda_j)}. \quad (39)$$

In this case $\lambda_1 = e^{p_f}$ corresponds, in the WKB-language, to the $p_f < 0$ crossing point of the activation trajectory and the p -axis, see Fig. 2. Therefore, to set the ground for matching Eq. (38) with the WKB solution in scenario A, we can rewrite Eq. (38) as

$$\pi(q) \simeq \frac{A_1}{\alpha\tau N} \frac{e^{-Np_f q}}{q}. \quad (40)$$

In scenario B the $n \gg 1$ asymptote of Eq. (37) is quite different. Here the root of Eq. (31) with the smallest absolute value is $\lambda_0 = 1$, see Appendix C. Therefore, the $i = 0$ term in Eq. (37) is dominant, and we obtain

$$\pi(q) \simeq \frac{1}{\alpha\tau Nq [1 - w'_1(0) - 2w'_2(0) - \dots - Kw'_K(0)]}, \quad (41)$$

where we have used Eqs. (35) and (36). The asymptote (41) can be expressed in terms of the WKB Hamiltonian (12). Using Eq. (16), we obtain $H_{qp}(0, 0) =$

$\sum_r r w'_r(0)$. Recalling that $w'_{-1}(0) = 1$ and using Eq. (29), we can rewrite $H_{qp}(0, 0)$ as

$$H_{qp}(0, 0) = K w'_K(0) + (K-1) w'_{K-1}(0) + \dots + w'_1(0) - 1. \quad (42)$$

As $q = 0$ is an attracting point here, $H_{qp}(0, 0) < 0$. Then, using Eq. (42), the asymptote (41) becomes

$$\pi(q) \simeq \frac{1}{\alpha \tau N |H_{qp}(0, 0)| q}. \quad (43)$$

Note that $\lambda_0 = 1 = e^{p_s}$ corresponds to the zero-momentum ($p_s = 0$) crossing point of the relaxation trajectory and the p -axis, see Fig. 3.

IV. EXTINCTION SCENARIO A

A. General case: multi-step processes

In this section we calculate the MTE and QSD for extinction scenario A. Here extinction occurs along the activation trajectory: the heteroclinic trajectory, connecting the metastable point $(n_1, 0)$ and the fluctuational extinction point $(0, p_f)$ of the phase plane (n, p) , see Fig. 2. In this case the slow-mode solution is negligible compared to the fast-mode solution in the entire region of $q > 0$. Furthermore, the fast-mode solution (24) can be directly matched with the recursive solution (37) in the joint region of their validity which turns out to be $1 \ll n \ll N^{1/2}$, or $N^{-1} \ll q \ll N^{-1/2}$.

To implement the matching procedure, we first find the $q \ll N^{-1/2}$ asymptote of the fast-mode solution (24). Because of the divergence of $S_1^{(f)}(q)$ at $q = 0$, we should proceed with care. Let us rewrite Eq. (24) as

$$\pi(q) = \frac{\sqrt{S''(q_1)}}{\sqrt{2\pi N} q} e^{N[S^{(f)}(q_1) - S^{(f)}(q)] + S_1^{(f)}(q_1) - [S_1^{(f)}(q) - \ln q]}. \quad (44)$$

Here we have introduced the $1/q$ prefactor which diverges at $q = 0$, and made up for it by adding $\ln q$ in the exponent. Let us show that the expression $S_1^{(f)}(q) - \ln q$ in the exponent is regular at $q = 0$. We represent $\ln q$ as $\int^q d\xi/\xi$ and use Eq. (17) to rewrite $S_1^{(f)}(q) - \ln q$ as an integral over ξ . Now we Taylor-expand the integrand

$$\frac{H_{pq}(\xi, p_a) + \frac{1}{2} H_{pp}(\xi, p_a) p'_a(\xi) - \sum_r u_r(\xi) (e^{r p_a} - 1)}{H_p(\xi, p_a)} - \frac{1}{\xi}$$

in the vicinity of $\xi = 0$ up to linear terms. The divergent terms cancel out, and the remaining expression

$$\frac{(1/2) H_{ppq}(0, p_f) p'_a(0) - \sum_r u'_r(0) (e^{r p_f} - 1)}{H_{pq}(0, p_f)}, \quad (45)$$

is finite. Now we rewrite Eq. (44) as

$$\pi(q) = \frac{\sqrt{S''(q_1)}}{\sqrt{2\pi N}} \frac{q_1}{q} e^{N[S^{(f)}(q_1) - S^{(f)}(q)] + \phi(q_1) - \phi(q)}, \quad (46)$$

where

$$\phi(q) = S_1^{(f)}(q) - \ln q \quad (47)$$

is regular at $q = 0$. By Taylor-expanding the exponent of Eq. (46) around $q = 0$ to first order, we obtain the $q \ll N^{-1/2}$ asymptote of the fast-mode solution:

$$\pi(q) \simeq \frac{\sqrt{S''(q_1)}}{\sqrt{2\pi N}} \frac{q_1}{q} e^{-N q p_f} e^{N[S^{(f)}(q_1) - S^{(f)}(0)] + \phi(q_1) - \phi(0)}. \quad (48)$$

This asymptote can be matched with the asymptote of the recursive solution at $N^{-1} \ll q \ll N^{-1/2}$, given by Eq. (40). This matching yields

$$\tau = \frac{A_1 \sqrt{2\pi}}{\alpha q_1 \sqrt{N S''(q_1)}} e^{N[S^{(f)}(0) - S^{(f)}(q_1)] + \phi(0) - \phi(q_1)}, \quad (49)$$

where $\phi(q)$ is given by Eq. (47), α is the linear decay rate constant in physical units, and A_1 is given by Eq. (39) [26]. The general expression (49) for the MTE in scenario A is one of the main results of this work. The leading term in the exponent, proportional to N , is the effective entropy barrier to extinction. The proportionality factor is the absolute value of the area under the activation trajectory, see an example in Fig. 2 [13]. Noticeable is the presence of the large factor $N^{1/2}$ in the pre-exponent. The constant A_1 has a clearly non-WKB nature, as it comes from the recursive solution of the quasi-stationary master equation at small n and is contributed to by *all* of the roots λ_i , $i = 0, 1, \dots, K$.

Another important result is the QSD in extinction scenario A. It is determined by the asymptotes (24) and (37) which coincide, in the leading order, in their joint region of validity $N^{-1} \ll q \ll N^{-1/2}$, or $1 \ll n \ll N^{1/2}$.

B. Single-step processes

Remaining within scenario A, we now turn to an important sub-class of stochastic population processes: single-step processes. Here there are only two non-zero process rates: $W_{\pm 1}(Nq) \equiv W_{\pm}(Nq) = N w_{\pm}(q) + u_{\pm}(q) + \dots$, where all the rates are normalized by the linear decay rate constant $w'_-(0)$. In this case the expressions for the MTE and QSD can be simplified considerably. The WKB Hamiltonian (12) becomes

$$H(q, p) = w_+(q)(e^p - 1) + w_-(q)(e^{-p} - 1). \quad (50)$$

The rate equation is $\dot{q} = w_+(q) - w_-(q)$. In scenario A one has $w'_+(0) > w'_-(0) = 1$. Here it is convenient to denote the ratio of the linear birth and death rates by $R \equiv w'_+(0)/w'_-(0) = w'_+(0)$. For $R > 1$ the fixed point $q = 0$ of the rate equation is repelling. The activation trajectory is $p_a(q) = -\ln[w_+(q)/w_-(q)]$, and

$$S^{(f)}(q) = - \int^q \ln \frac{w_+(\xi)}{w_-(\xi)} d\xi. \quad (51)$$

Now we calculate the following quantities on the activation trajectory:

$$\begin{aligned} p'_a(q) &= S''(q) = \frac{w'_-}{w_-} - \frac{w'_+}{w_+}, \quad H_p(q, p_a) = w_- - w_+, \\ H_{pq}(q, p_a) &= \frac{w_- w'_+}{w_+} - \frac{w_+ w'_-}{w_-}, \quad H_{pp}(q, p_a) = w_- + w_+, \\ \sum_r u_r(e^{r p_a} - 1) &= u_+ \left(\frac{w_-}{w_+} - 1 \right) + u_- \left(\frac{w_+}{w_-} - 1 \right). \end{aligned} \quad (52)$$

Substituting these into Eq. (17), we obtain after simplifications

$$S_1^{(f)}(q) = \int^q \left(\frac{u_-}{w_-} - \frac{u_+}{w_+} \right) d\xi + \frac{1}{2} \ln[w_+(q)w_-(q)]. \quad (53)$$

Plugging this into Eq. (47) yields

$$\phi(q) = \int^q \left(\frac{u_-}{w_-} - \frac{u_+}{w_+} \right) d\xi + \frac{1}{2} \ln \left[\frac{w_+(q)w_-(q)}{q^2} \right]. \quad (54)$$

Now we can calculate $e^{\phi(0)-\phi(q_1)}$ which enters Eqs. (48) and (49):

$$e^{\phi(0)-\phi(q_1)} = \frac{q_1 \sqrt{R}}{w_+(q_1)} \exp \left[\int_0^{q_1} \left(\frac{u_+}{w_+} - \frac{u_-}{w_-} \right) dq \right], \quad (55)$$

where we have used the following relations: (i) as $q \rightarrow 0$, $w_+(q)w_-(q)/q^2 \rightarrow w'_+(0)w'_-(0) = R$, (ii) $w_+(q_1) = w_-(q_1)$ (as q_1 is a fixed point of the rate equation), and (iii) $w'_-(0) = 1$ because of the rescaling of the rates.

Now we turn to the recursive solution at small n , presented in Section III. For single-step processes $K = 1$, and so Eq. (31) has only two roots: $\lambda_0 = 1$ and $\lambda_1 = 1/w'_+(0) = 1/R$. Therefore, rewriting the small- q asymptote (37) of the QSD in terms of π_n , we obtain

$$\pi_n = \frac{(R^n - 1)f_1}{(R - 1)n}, \quad (56)$$

whereas the constant A_1 from Eq. (39) is $1/(R-1)$. Plugging this constant and Eqs. (51) and (55) in Eq. (49) for the MTE, we obtain

$$\tau = \frac{\sqrt{2\pi R} e^{\int_0^{q_1} \left(\frac{u_+}{w_+} - \frac{u_-}{w_-} \right) dq}}{\alpha(R-1)w_+(q_1)\sqrt{NS''(q_1)}} e^{N \int_0^{q_1} \ln \left(\frac{w_+}{w_-} \right) dq}. \quad (57)$$

In the particular case $u_+ = u_- = 0$ Eq. (57) coincides with Eq. (19) of Ref. [18], obtained, via a saddle-point approximation, from the exact expression for the MTE of a single-step process. Doering *et al.* assumed in their derivation that the subleading contributions u_{\pm} to the process rates $W_{\pm 1}(n)$, see Eq. (9), vanish. As a result, the factor $\exp[\int_0^{q_1} (u_+/w_+ - u_-/w_-) dq]$ is absent from their Eq. (19). While the assumption $u_+ = u_- = 0$ may hold in some simple models, it does not hold in general. For example, it does not hold for stochastic chemical reactions where the rates are combinatorial, as in one of the examples we present in subsection IVD below.

C. Extinction near transcritical bifurcation point

Now let us return to a general set of (not necessarily single-step) processes. Our objective is to simplify the MTE (49) in the special regime when the population, as described by the rate equation, is very close to the characteristic (transcritical) bifurcation point of scenario A. Here the attracting point $q = q_1$ is very close to the repelling point $q = 0$, so that $q_1 \ll 1$. This also implies $|p_f| \ll 1$ [14, 27]. Taylor-expanding Eq. (12) in q and p around $q = p = 0$, we obtain

$$H(q, p) \simeq qp \sum_r \left[r w'_r(0) + \frac{q}{2} r w''_r(0) + \frac{p}{2} r^2 w'_r(0) \right] = 0. \quad (58)$$

The trivial solutions are the extinction line $q = 0$ and the relaxation trajectory $p = 0$, whereas the nontrivial solution yields a straight-line activation trajectory. Using Eq. (16) and expanding the algebraic equations $H_p(q, 0) = 0$ for q_1 , and $H_q(0, p) = 0$ for p_f at small q and p , we can represent the activation trajectory as

$$p_a(q) = -p_f \left(\frac{q}{q_1} - 1 \right). \quad (59)$$

Here

$$q_1 = -\frac{2H_{qp}(0, 0)}{H_{qqp}(0, 0)}, \quad (60)$$

where $H_{qqp}(0, 0) < 0$, and

$$p_f = -\frac{2H_{qp}(0, 0)}{H_{qpp}(0, 0)}, \quad (61)$$

where $H_{qpp}(0, 0) > 0$. Exactly at the bifurcation the rate constants are such that $H_{qp}(0, 0) = 0$. Here the attracting fixed point q_1 merges with the repelling point $q = 0$. The coordinate of the attracting fixed point $q_1 \equiv \delta$ can serve here as the distance to the bifurcation. [The third derivatives of the Hamiltonian, which appear in the denominators of Eqs. (60) and (61), are generically of order unity.]

Now, using Eq. (59), we can calculate $S''(q_1) = p'_a(q_1) = -p_f/q_1 = -p_f/\delta > 0$, the fast-mode action

$$S^{(f)}(q) = \int^q p_a(\xi) d\xi = p_f q - \frac{p_f}{\delta} \frac{q^2}{2}, \quad (62)$$

and the accumulated action between the points $q = q_1 = \delta$ and $q = 0$

$$\Delta S = S^{(f)}(0) - S^{(f)}(q_1) = -\frac{p_f q_1}{2} = -\frac{p_f \delta}{2} > 0. \quad (63)$$

This quantity is the area of a triangle [14], see Fig. 4.

To find the fast-mode correction to the action, $S_1^{(f)}(q)$, we expand Eq. (15) in the vicinity of $q = 0$ and $p = 0$, keeping only the leading order terms. This yields

$$[q S'_1(q) - 1] \sum_r r w'_r(0) \simeq 0, \quad (64)$$

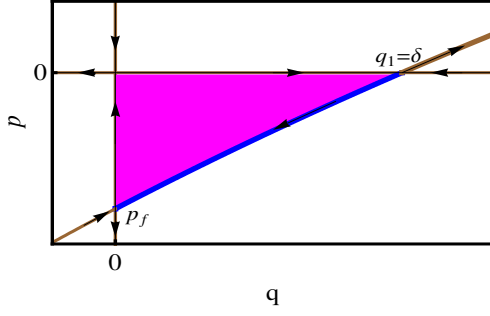


FIG. 4: (color online). Shown are typical zero-energy trajectories of the WKB Hamiltonian (12) in scenario A, close to the bifurcation where $q_1 = \delta \ll 1$, and $p_f \ll 1$. Here, the activation trajectory is a straight line, and ΔS given by Eq. (63) is the area of the shaded triangle.

whereas the subleading terms u_r in the rate expansion do not contribute. The solution of this equation is $S_1^{(f)}(q) = \ln q$. Then, by virtue of Eq. (47), we obtain $\phi(q) = 0$. That is, near the bifurcation, the subleading WKB correction vanishes.

Now let us consider the coefficient A_1 [see Eq. (39)] which enters Eq. (49). Among the roots λ_i of the polynomial (31), which contribute to A_1 , two are special: $\lambda_0 = 1$ and $\lambda_1 = e^{p_f}$. Near the bifurcation $|p_f| \ll 1$, so we can write $\lambda_1 \simeq 1 + p_f < 1$. As a result, $A_1 = \tilde{A}_1/p_f$, where

$$\tilde{A}_1 = \frac{(-1)^K \prod_{j=2}^K \lambda_j}{\prod_{j=2}^K (1 - \lambda_j)} \quad (65)$$

is a negative constant of order unity, and we have put $\lambda_1 = 1$ in the expression for \tilde{A}_1 . Furthermore, the roots $\lambda_2, \lambda_3, \dots, \lambda_K$ of the polynomial (33) can be evaluated at the bifurcation point. As a result, one can express \tilde{A}_1 via the linear branching rates $w'_r(0)$. After some algebra

$$\tilde{A}_1 = -2 \left[\sum_{r=1}^K r(r+1)w'_r(0) \right]^{-1}. \quad (66)$$

Substituting all of the above into Eq. (49), we obtain the MTE close to the bifurcation point:

$$\tau = \sqrt{\frac{2\pi}{N}} \frac{|\tilde{A}_1|}{\alpha D_A^3 \delta^2} \exp\left(\frac{N D_A^2 \delta^2}{2}\right), \quad (67)$$

where \tilde{A}_1 and $D_A = \sqrt{|H_{qpp}(0,0)|/H_{qpp}(0,0)} = \mathcal{O}(1)$ should be evaluated at the bifurcation. Equation (67) is valid when $N\delta^2 \gg 1$. For sufficiently large N this strong inequality is compatible with the strong inequality $\delta \ll 1$ which describes closeness to the bifurcation. Note that the constant $\tilde{A}_1 = \mathcal{O}(1)$ is determined by the full small- n recursive solution that we found in section III.

Although Eq. (79) breaks down at $\delta \sim N^{-1/2}$, one can still predict a scaling relation for the MTE in this region: $\tau \sim N^{1/2}/\alpha$. The symbol \sim , here and in the following, means “of the same order as”.

D. Examples

We will now illustrate our theory by calculating the MTE in four pedagogical examples of extinction scenario A. The first three of them are single-step processes: the logistic Verhulst model of population dynamics, a set of three chemical reactions, and the SIS model of epidemics. The fourth example - another set of three chemical reactions - involves a two-step process. We will also consider all of these examples near the bifurcation.

1. Verhulst model

The generalized Verhulst model is a stochastic logistic model: a single-step Markov process with birth and death rates

$$\begin{aligned} W_{+1} &\equiv W_+ = \alpha_1 n - \alpha_2 n^2 \\ W_{-1} &\equiv W_- = \beta_1 n + \beta_2 n^2, \end{aligned} \quad (68)$$

respectively, where $\alpha_1, \alpha_2, \beta_1$ and β_2 are non-negative rate constants. The quadratic corrections account for competition for resources [28]. It is customary to put $\alpha_2 = 0$ in Eq. (68) [18], and this is what we will do here. Rescaling time by the linear death rate constant β_1 , we bring the rates to the form given by Eq. (9): $W_+ = Bn$ and $W_- = n + Bn^2/N$, where $B = \alpha_1/\beta_1$ is the ratio of the linear birth and death rates, and $N = B\beta_1/\beta_2$. According to Eq. (9) $w_+ = Bq$, $w_- = q + Bq^2$, and $u_+ = u_- = 0$. At $B > 1$ the fixed point $q = 0$ of the rate equation is repelling, whereas $q_1 = 1 - 1/B > 0$ is attracting. Here we have $S''(q_1) = p'_a(q_1) = 1$, $A_1 = 1/(B - 1)$ and

$$\int_0^{q_1} \ln \frac{w_-}{w_+} dq = \frac{1 - B + \ln B}{B}.$$

Therefore, the MTE [Eq. (57)] in physical time units is

$$\tau = \frac{1}{\beta_1} \sqrt{\frac{2\pi}{N}} \frac{\sqrt{B}}{(B-1)^2} \exp\left[N \left(\frac{B-1-\ln B}{B}\right)\right], \quad (69)$$

which coincides with previous results obtained by different methods [18, 28].

In this simple example the process rates satisfied the conditions $u_+ = u_- = 0$, so Eq. (69) could have been obtained from Eq. (19) of Ref. [18]. In the next example we relax one of the two conditions and show that, as predicted by our more general Eq. (57), the pre-exponent of the MTE changes.

2. Reactions $A \rightleftharpoons 2A$ and $A \rightarrow 0$

Consider a set of three reactions among particles A : branching $A \xrightarrow{\lambda} 2A$, a reverse reaction $2A \xrightarrow{\sigma} A$, and decay $A \xrightarrow{\mu} \emptyset$. As observed in Ref. [27], this set of reactions can be viewed as a generalization of the Verhulst model considered in the previous example. Indeed, by imposing a special relation, $\sigma = 2(\mu - 1)$, between the rate constants, and by denoting $\mu = 1 + B/N$ and $\lambda = B$, one recovers the process rates w_{\pm} and u_{\pm} of the Verhulst model. For this special choice of rate constants one has $u_{+} = u_{-} = 0$ which yields Eq. (69) for the MTE. Now, what if the rate constants σ and μ are independent? As usual, we can normalize time and the reaction rates by μ and denote $\tilde{B} = \lambda/\mu$ and $\tilde{N} = 2\lambda/\sigma$. By virtue of Eq. (9) we can write $w_{+} = \tilde{B}q$, $w_{-} = q + \tilde{B}q^2$, $u_{+} = 0$ and $u_{-} = \tilde{B}q$. The rescaled rate constants are identical to those of the Verhulst model, except that u_{-} is now nonzero. As a result,

$$\exp \left[\int_0^{q_1} \left(\frac{u_{+}}{w_{+}} - \frac{u_{-}}{w_{-}} \right) dq \right] = \tilde{B},$$

and Eq. (57) for the MTE yields

$$\tau = \sqrt{\frac{\pi\sigma}{\mu}} \frac{\lambda}{(\lambda - \mu)^2} \exp \left[\frac{2}{\sigma} \left(\lambda - \mu + \mu \ln \frac{\mu}{\lambda} \right) \right], \quad (70)$$

where we have returned to physical units. This result cannot be obtained from Eq. (19) of Ref. [18].

3. SIS model

Now let us consider the well-known SIS model of epidemics, see *e.g.* Refs. [28, 29] and references therein. The SIS model deals with dynamics of a population which consists of two groups of individuals: susceptible to infection and infected. It is assumed that infection does not confer any long-lasting immunity, and infected individuals become susceptible again after infection. When demography (births and deaths) is negligible, the total number N of individuals in the two groups is conserved. As a result, the model becomes effectively single-population, with the effective rates

$$W_{+} = \lambda n(N - n), \quad W_{-} = \mu n.$$

Mathematically, this model is just another example of the generalized Verhulst model, see Eq. (68), where one chooses $\beta_1 \neq 0$ but $\beta_2 = 0$ [28].

Let us denote $R_0 = \lambda N/\mu$ and rescale time and rates by the linear decay rate constant $\mu = \lambda N/R_0$. The rescaled rates become $w_{+} = R_0(q - q^2)$ and $w_{-} = q$, while $u_{+} = u_{-} = 0$. The fixed point $q_1 = 1 - 1/R_0$ of the rate equation is attracting when $R_0 > 1$. Furthermore, $S''(q_1) = p'_a(q_1) = R_0$, and $A_1 = 1/(R_0 - 1)$ (as in the above notation $R = R_0$). Finally,

$$\int_0^{q_1} \ln \frac{w_{-}}{w_{+}} dq = 1 - \frac{1}{R_0} - \ln R_0.$$

Therefore, the MTE [Eq. (57)], in physical time units, is given by

$$\tau = \frac{1}{\mu} \sqrt{\frac{2\pi}{N}} \frac{R_0}{(R_0 - 1)^2} \exp \left[N \left(\ln R_0 + \frac{1}{R_0} - 1 \right) \right], \quad (71)$$

which coincides with previous results obtained by different methods [28–30].

4. Branching-annihilation-decay

Now we consider another set of stochastic reactions among particles A which include, in addition to single-step processes $A \xrightarrow{\lambda} 2A$ and $A \xrightarrow{\mu} \emptyset$, a two-step process: binary annihilation $2A \xrightarrow{\sigma} \emptyset$. This problem was previously solved by Kessler and Shnerb [16]. Here we show that their result for the MTE follows from our Eq. (49).

In our notation, the transition rates between the states n and $n + r$ are given by

$$W_1 = \lambda n, \quad W_{-1} = \mu n, \quad \text{and} \quad W_{-2} = \frac{\sigma n(n-1)}{2}. \quad (72)$$

Rescaling time $\mu t \rightarrow t$ and denoting $R_0 = \lambda/\mu$ and $N = \lambda/\sigma$, we obtain Eq. (9) with

$$\begin{aligned} w_1 &= R_0 q, & w_{-1} &= q, & w_{-2} &= \frac{R_0 q^2}{2}, \\ u_1 &= 0, & u_{-1} &= 0, & u_{-2} &= -\frac{R_0 q}{2}. \end{aligned} \quad (73)$$

In the rescaled notation, the attracting fixed point is $q_1 = 1 - 1/R_0$ which demands $R_0 > 1$. The WKB Hamiltonian (12) takes the form

$$H(q, p) = R_0 q (e^p - 1) + q (e^{-p} - 1) + \frac{R_0 q^2}{2} (e^{-2p} - 1). \quad (74)$$

Solving the equation $H[q, p_a(q)] = 0$, we obtain the activation trajectory

$$p_a(q) = S'(q) = \ln \left(\frac{u + v}{4R_0} \right), \quad (75)$$

where $u = 2 + qR_0$ and $v = \sqrt{u^2 + 8qR_0^2}$. The zero-energy phase trajectories of this system are shown in Fig. 2.

Now we use Eqs. (14), (17), and (47) and obtain

$$\begin{aligned} \Delta S &= S^{(f)}(0) - S^{(f)}(q_1) \\ &= 2 \left\{ 1 - \frac{1}{R_0} + \left(1 + \frac{1}{R_0} \right) \ln \left[\frac{1}{2} \left(1 + \frac{1}{R_0} \right) \right] \right\} \end{aligned} \quad (76)$$

and

$$e^{\phi(0) - \phi(q_1)} = \frac{2R_0}{\sqrt{(R_0 + 1)(3R_0 - 1)}}. \quad (77)$$

Furthermore, $S''(q_1) = 2R_0/(3R_0 - 1)$ and, as in the previous examples, the only root of Eq. (33) is $\lambda_1 = 1/R_0$.

Therefore, by using Eq. (39), we have $A_1 = 1/(R_0 - 1)$. Substituting all of the above into Eq. (49), we obtain, in physical time units,

$$\tau = \frac{2\sqrt{\pi}}{\mu\sqrt{N}} \frac{R_0^{3/2}}{(R_0 - 1)^2(R_0 + 1)^{1/2}} e^{N\Delta S} \quad (78)$$

which coincides with the result of Ref. [16].

5. Examples 1-4 near the bifurcation point

Because of their simplicity the examples 1-4, presented above, give identical results for the MTE near their corresponding bifurcation points, described by the equation $H_{qp}(0,0) = 0$. The small distance to the bifurcation δ in all these examples is the ratio of the linear birth and death rates minus 1: $B - 1$ in example 1, $\lambda/\mu - 1$ in example 2, and $R_0 - 1$ in examples 3 and 4. In all four examples $D_A = \sqrt{|H_{qqp}(0,0)|/H_{qpp}(0,0)} = 1$. As there is only one linear branching process in each example, one has $K = 1$, so Eq. (66) yields $\tilde{A}_1 = -1$. As a result, in all four examples

$$\tau = \frac{\sqrt{2\pi}}{\alpha\sqrt{N}\delta^2} \exp\left(\frac{N\delta^2}{2}\right), \quad (79)$$

where α denotes, in each example, the linear decay rate constant. To remind the reader, Eq. (79) is valid when $N\delta^2 \gg 1$ which, together with $\delta \ll 1$, yields the double inequality $N^{-1/2} \ll \delta \ll 1$. At $\delta \sim N^{-1/2}$ Eq. (79) predicts the following scaling relation for the MTE: $\tau \sim N^{1/2}/\alpha$, in agreement with Ref. [28].

V. EXTINCTION SCENARIO B

A. General case: multi-step processes

In this section we calculate the MTE and QSD for extinction scenario B. Here extinction occurs along a trajectory composed of two segments: the non-zero-momentum heteroclinic trajectory connecting the hyperbolic fixed points $(n_2, 0)$ and $(n_1, 0)$ (the activation trajectory), and the zero-momentum segment going from $n = n_1$ to $n = 0$ (the relaxation trajectory), see Fig. 3.

A straightforward way to calculate the MTE starts with finding the WKB solution for the QSD at $n \gg 1$. Then one should match it with the small- n recursive solution (37), as in scenario A. The matching region in this case is $1 \ll n \ll N$, or $N^{-1} \ll q \ll 1$. After having found the QSD, one can determine the MTE by using Eq. (43).

Actually, there is a shortcut to finding the MTE which does not require the knowledge of the small- n recursive solution. This is because the solution includes a constant probability current flowing from a close vicinity of $n = n_1$ to a close vicinity of $n = 0$, as shown below. This

probability current is equal to the escape rate from the metastable state $q = q_2$, and it is determined by the WKB-asymptote of the QSD, with no use of the small- n recursive solution. One of the objectives of our work, however, is to also find the QSD of the metastable state, and therefore we will follow the straightforward way.

In contrast to scenario A, where the WKB solution is determined solely by the fast mode, in scenario B the WKB solution is more complicated. Here the fast mode dominates to the right of the point $q = q_1$ (but not too close to q_1), whereas the slow-mode solution (26) dominates at $0 < q < q_1$ (again, not too close to $q = q_1$). Furthermore, the slow-mode solution diverges at $q = q_1$, and curing this divergence demands going beyond the WKB approximation in a boundary layer $|q - q_1| \ll 1$ where the fast and slow modes are strongly coupled. As a result, the QSD at $n \gg 1$ involves *three* distinct asymptotes which need to be matched to one another. All this is very similar to what happens in other types of population escape problems: to an absorbing state at infinite population size [21] or to another metastable state [22]. Much of the calculation is very similar to that of Refs. [21, 22], but we will present it here for completeness.

In the boundary layer $|q - q_1| \ll 1$ the momentum p is small, that is fluctuations are weak. Here we can apply the van Kampen system size expansion [5] to the quasi-stationary master equation (7). Let us denote $f(q) = W_r(q)\pi(q) \simeq Nw_r(q)\pi(q)$ [it suffices to keep only the leading term in Eq. (9)]. Taylor-expanding $f(q - r/N)$ around $r = 0$, we obtain

$$f(q - r/N) \simeq f(q) - \frac{r}{N}f'(q) + \frac{r^2}{2N^2}f''(q). \quad (80)$$

Plugging Eq. (80) into Eq. (7) and integrating once, we obtain

$$\sum_r -\frac{r}{N}f(q) + \frac{r^2}{2N^2}f'(q) = \tilde{J}, \quad (81)$$

where $\tilde{J} = \text{const.}$ Now,

$$f'(q) = N[\pi'(q)w_r(q) + \pi(q)w'_r(q)],$$

but $\pi'(q) \sim N\pi(q)$, so the second term in $f'(q)$ is negligible. Therefore, we obtain

$$-\pi(q) \sum_r rw_r(q) + \pi'(q) \sum_r \frac{r^2}{2N}w_r(q) = \tilde{J}. \quad (82)$$

The first term on the left corresponds to drift, the second one to diffusion. With the diffusion neglected, one obtains a (slow-mode) solution for $\pi(q)$ which diverges at fixed points of the rate equation. The diffusion term cures this divergence by providing coupling between the slow mode and fast modes, as observed in Ref. [21].

Now we use Eq. (16) and evaluate the drift and diffusion terms in the vicinity of $q = q_1$. In the drift term

$$\sum_r rw_r(q) = H_p(q, 0) \simeq (q - q_1)H_{pq}(q_1, 0),$$

while in the diffusion term it suffices to put $w_r(q) \simeq w_r(q_1)$. Denoting $x = (q - q_1)/l$, and $l^2 = H_{pp}(q_1, 0)/[NH_{pq}(q_1, 0)]$ [as one can check, both $H_{pp}(q_1, 0)$ and $H_{pq}(q_1, 0)$ are positive], we obtain the boundary-layer equation [21]

$$\pi'(x) - 2x\pi(x) = J, \quad (83)$$

where the rescaled constant current J is to be found later. The general solution of Eq. (83) is

$$\pi(x) = c_1 e^{x^2} + \frac{\pi J}{2} \text{erf}(x). \quad (84)$$

where c_1 is another constant. Now we can match this solution to the slow-mode solution at $x < 0$ and $|x| \gg 1$, that is, at $N^{-1/2} \ll q_1 - q \ll 1$. To eliminate the exponential growth at $x < 0$, one must choose $c_1 = J\sqrt{\pi}/2$, so the asymptote of the boundary-layer solution (84) at $-x \gg 1$ becomes

$$\pi(x) \simeq -\frac{J}{2x} = \frac{J}{2(q_1 - q)} \sqrt{\frac{H_{pp}(q_1, 0)}{NH_{pq}(q_1, 0)}}. \quad (85)$$

The slow-mode solution (26) at $q \simeq q_1$ can be approximated as

$$\pi_s(q) = -\frac{A_s}{H_p(q, 0)} \simeq \frac{A_s}{(q_1 - q)H_{pq}(q_1, 0)}. \quad (86)$$

Matching the two asymptotes, one obtains

$$J = \frac{2A_s\sqrt{\pi N}}{\sqrt{H_{pp}(q_1, 0)H_{pq}(q_1, 0)}}. \quad (87)$$

To find the still unknown constant A_s , we have to match the $x \gg 1$ asymptote of the boundary-layer solution (84), which is

$$\begin{aligned} \pi(x) &\simeq J\sqrt{\pi}e^{x^2} = \frac{2A_s\sqrt{\pi N}}{\sqrt{H_{pp}(q_1, 0)H_{pq}(q_1, 0)}} \\ &\times \exp\left[\frac{NH_{pq}(q_1, 0)}{H_{pp}(q_1, 0)}(q - q_1)^2\right], \end{aligned} \quad (88)$$

with the asymptote of the fast-mode solution at $N^{-1/2} \ll q - q_1 \ll 1$, which is

$$\begin{aligned} \pi(q) &\simeq \sqrt{\frac{S''(q_2)}{2\pi N}} \\ &\times e^{N[S^{(f)}(q_2) - S^{(f)}(q_1)] + S_1^{(f)}(q_2) - S_1^{(f)}(q_1) - (N/2)S''(q_1)(q - q_1)^2}. \end{aligned} \quad (89)$$

Here we have used the equalities $S'(q_1) = p_a(q_1) = 0$ and neglected terms of order $q - q_1 \ll 1$ in the exponent. Putting $q = q_1$ into Eq. (19), we obtain

$$p'_a(q_1) = S''(q_1) = -\frac{2H_{pq}(q_1, 0)}{H_{pp}(q_1, 0)}, \quad (90)$$

where $S''(q_1) < 0$. Matching the asymptotes (88) and (89) and using Eq. (90), we find

$$\begin{aligned} A_s &= \frac{H_{pp}(q_1, 0)\sqrt{|S''(q_1)|S''(q_2)}}{4\pi N} \\ &\times e^{N[S^{(f)}(q_2) - S^{(f)}(q_1)] + [S_1^{(f)}(q_2) - S_1^{(f)}(q_1)]}. \end{aligned} \quad (91)$$

What is left is to find the MTE by matching the slow-mode solution at $q \ll 1$ with the recursive solution (43) at $q \gg N^{-1}$. Using Eq. (26) we obtain, at $q \ll 1$:

$$\pi_s(q) \simeq -\frac{A_s}{qH_{pq}(0, 0)} = \frac{A_s}{q|H_{pq}(0, 0)|}, \quad (92)$$

where $H_{pq}(0, 0) < 0$ is given by Eq. (42). Comparing this with Eq. (43) and using Eq. (91), we obtain

$$\begin{aligned} \tau &= (\alpha N A_s)^{-1} = \frac{4\pi}{\alpha H_{pp}(q_1, 0)\sqrt{|S''(q_1)|S''(q_2)}} \\ &\times e^{N[S^{(f)}(q_1) - S^{(f)}(q_2)] + [S_1^{(f)}(q_1) - S_1^{(f)}(q_2)]}, \end{aligned} \quad (93)$$

where α in the linear decay rate constant in physical units [26]. The expression (93) for the MTE in scenario B is an important result of our work. The leading term in the exponent, proportional to N , is the effective entropy barrier to extinction. The proportionality factor is the absolute value of the area between the activation trajectory and relaxation trajectory [14], see an example in Fig. 3. In contrast to scenario A, the pre-exponential factors in Eq. (93) are N -independent. Using Eqs. (20), (21) and (90), one can rewrite Eq. (93) in a more concise form:

$$\tau = \frac{2\pi}{\alpha H_{pq}(q_1, 0)} e^{N[S^{(f)}(q_1) - S^{(f)}(q_2)] + [\Psi(q_1) - \Psi(q_2)]}. \quad (94)$$

As mentioned above, determining the MTE in scenario B does not require any information about the small- n recursive solution. Furthermore, Eq. (93) formally coincides with the result of Escudero and Kamenev [22], who calculated a different quantity: the mean time to *escape* from one metastable state into another. Finally, the same result (93) can be also obtained for the mean time to escape to an absorbing state at infinity, as in the particular example considered by Meerson and Sasorov [21]. The reason for these coincidences is that, in all these systems, a constant probability current sets in beyond the repelling fixed point of the rate equation. It is the magnitude of this current, carried by the slow WKB mode, rather than the exact nature of the target state for escape (an absorbing state at zero, infinity or another metastable state), that determines, in the leading and subleading orders in N , the mean escape rate from a metastable state.

To conclude this section, the QSD (another main result of this work) is given by four overlapping asymptotes: (i) the recursive solution (37), valid for $1 \leq n \ll N$, (ii) the slow-mode WKB solution (26), valid for $n_1 - n \gg N^{1/2}$ and $n \gg 1$, (iii) the boundary-layer solution (84), valid for $|n - n_1| \ll n_1$, and (iv) the fast-mode WKB solution (24), valid for $n - n_1 \gg N^{1/2}$.

B. Single-step processes

For completeness, we briefly consider the special case of single-step processes, where only $W_{\pm 1}(Nq) \equiv W_{\pm}(Nq) = Nw_{\pm}(q) + u_{\pm}(q) + \dots$ are present. Here Eq. (93) simplifies considerably. Performing calculations similar to those in scenario A (see Sec. IV) and using Eqs. (16) and (53) and the fact that $w_+(q_{1,2}) = w_-(q_{1,2})$, one obtains the MTE

$$\tau = \frac{2\pi e^{\int_{q_1}^{q_2} \left(\frac{u_+}{w_+} - \frac{u_-}{w_-}\right) dq}}{\alpha w_+(q_2) \sqrt{|S''(q_1)|S''(q_2)}} e^{N \int_{q_2}^{q_1} \ln\left(\frac{w_-}{w_+}\right) dq}. \quad (95)$$

As expected, this result coincides with the single-step result of Ref. [22] for the mean time of a population switch between two metastable states.

C. Extinction near saddle-node bifurcation point

Here we calculate the MTE near the characteristic (saddle-node) bifurcation of scenario B. At the bifurcation, the nontrivial attracting fixed point $q = q_2$ of the rate equation merges with the repelling point $q = q_1$. Above but near the bifurcation point $q_2 - q_1 \ll 1$. As a result, the momentum p on the activation trajectory is much smaller than unity, see Fig. 5. One can always define the parameter N such that, at the bifurcation, $q_1 = q_2 = 1$. Furthermore, near the bifurcation $q_1 = 1 - \delta$ and $q_2 = 1 + \delta$, where the exact definition of $\delta \ll 1$ will appear shortly.

Let us Taylor-expand $H(q, p)$ from Eq. (12) in the vicinity of $q = 1$ and $p = 0$. As we expect $p_a(q)$ to be $\sim (q_2 - q_1)^2$, we neglect the terms of order $(q - 1)p^2$ and higher and arrive at the following equation for the zero-energy phase trajectories $p = p(q)$ close to $q = 1$:

$$H(q, p) \simeq p \sum_r [rw_r(1) + (q - 1)rw'_r(1) + \frac{1}{2}(q - 1)^2rw''_r(1) + \frac{1}{2}pr^2w_r(1)] = 0. \quad (96)$$

As can be checked *a posteriori*, the terms in Eq. (96) scale as follows: $H_p(1, 0) \sim \delta^2$, $H_{qp}(1, 0) \sim \delta^2$, $H_{qqp}(1, 0) = \mathcal{O}(1)$, and $H_{pp}(1, 0) = \mathcal{O}(1)$. Therefore, the term $(q - 1)H_{qp}(1, 0) \sim \delta^3$ can be neglected. The nontrivial solution of Eq. (96) yields the activation trajectory $p = p_a(q)$: a parabola with the roots $q = q_1$ and $q = q_2$. To simplify the notation, we use Eq. (16) and evaluate the small difference $q_2 - q_1$ by expanding the algebraic equation $H_p(q, 0) = 0$ in the vicinity of $q = 1$. Neglecting the term $(q - 1)H_{qp}(1, 0)$, we obtain

$$\delta \equiv \frac{q_2 - q_1}{2} = \sqrt{\frac{2H_p(1, 0)}{|H_{qqp}(1, 0)|}}, \quad (97)$$

where $H_{qqp}(q_1, 0) < 0$. The activation trajectory can be written as

$$p_a(q) = \frac{|H_{qqp}(1, 0)|}{H_{pp}(1, 0)}(q - q_1)(q - q_2). \quad (98)$$

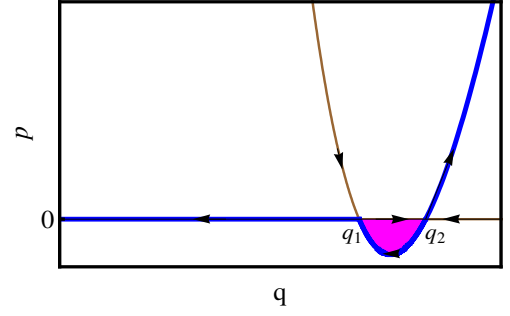


FIG. 5: (color online). Shown are typical zero-energy trajectories of the WKB Hamiltonian (12) in scenario B, close to the bifurcation where $q_2 - q_1 \ll 1$. Here, the activation trajectory is a parabola, and ΔS given by Eq. (100) is the area of the shaded region.

As $S''(q) = p'_a(q)$, we find $S''(q_2) = -S''(q_1) = 2\delta^2|H_{qqp}(1, 0)|/H_{pp}(1, 0)$. Furthermore, the action $S^{(f)}(q) = \int^q p_a(\xi)d\xi$ is given by

$$S^{(f)}(q) = \frac{|H_{qqp}(1, 0)|}{H_{pp}(1, 0)} \left[\frac{q^3}{3} - \frac{q^2}{2}(q_1 + q_2) + qq_1q_2 \right], \quad (99)$$

whereas

$$\Delta S = S^{(f)}(q_1) - S^{(f)}(q_2) = \frac{4|H_{qqp}(1, 0)|}{3H_{pp}(1, 0)}\delta^3 \quad (100)$$

is the area of the shaded region in Fig. 5.

As in scenario A, the sub-leading WKB correction vanishes near the bifurcation. Indeed, we Taylor-expand Eq. (15) in the vicinity of $q = 1$ and $p = 0$, keep only leading-order terms, and obtain

$$\sum_r rw_r(1)S'_1(q) - \frac{1}{2}r^2w_r(1)(q - 1)p''_a(1) - rw''_r(1) \simeq 0. \quad (101)$$

Using Eq. (98), we find that the second and third terms cancel out, and so $S'_1(q)$ can be chosen zero. As a result, Eq. (93) yields the MTE near the bifurcation:

$$\tau = \frac{2\pi}{\alpha|H_{qqp}(1, 0)|\delta} \exp\left(\frac{4}{3}ND_B^2\delta^3\right), \quad (102)$$

where $H_{qqp}(1, 0)$ and $D_B = \sqrt{|H_{qqp}(1, 0)|/H_{pp}(1, 0)} = \mathcal{O}(1)$ should be evaluated at the bifurcation. The applicability criterion of this result is $N\delta^3 \gg 1$. For sufficiently large N this strong inequality is compatible with the strong inequality $\delta \ll 1$ which describes closeness to the bifurcation. At $\delta \sim N^{-1/3}$ Eq. (102) predicts the following scaling of the MTE with N : $\tau \sim N^{1/3}/\alpha$.

We notice that Eq. (102) does not require any information about the QSD in the region of small n . Indeed, as was mentioned in section V A, the exact nature of the target state is of no significance here. Note that the same scaling of the effective entropy barrier with the distance from the bifurcation appears in the context of escape

from one metastable state to another [12]. Finally, the same scaling near the bifurcation is observed in *continuous* systems, driven by external delta-correlated gaussian noise and therefore describable by a Fokker-Planck equation [31, 32].

D. Example: reactions $2A \rightleftharpoons 3A$ and $A \rightarrow \emptyset$

Let us illustrate the extinction scenario B on the following set of reactions: binary reproduction $2A \xrightarrow{\lambda} 3A$, the reverse process $3A \xrightarrow{\sigma} 2A$, and linear decay $A \xrightarrow{\mu} \emptyset$. Here

$$W_1 = \frac{\lambda n(n-1)}{2}, \quad W_{-1} = \mu n + \frac{\sigma n(n-1)(n-2)}{6}. \quad (103)$$

Rescaling time $\mu t \rightarrow t$, and denoting $\gamma = 1 - \delta^2$, $\delta^2 = 1 - 8\sigma\mu/(3\lambda^2) > 0$, and $N = 3\lambda/(2\sigma)$, we arrive at Eq. (9) with

$$\begin{aligned} w_{-1} &= \frac{q^3}{\gamma} + q, & w_1 &= \frac{2q^2}{\gamma}, \\ u_{-1} &= -\frac{3q^2}{\gamma}, & u_1 &= -\frac{2q}{\gamma}. \end{aligned} \quad (104)$$

In the rescaled notation the fixed points are $q = 0$ (attracting point), $q_1 = 1 - \delta$ (repelling point), and $q_2 = 1 + \delta$: another attracting point around which the metastable population resides. The WKB Hamiltonian (12) takes the form

$$H(q, p) = \left(\frac{q^3}{\gamma} + q \right) (e^{-p} - 1) + \frac{2q^2}{\gamma} (e^p - 1). \quad (105)$$

Solving the equation $H[q, p_a(q)] = 0$ yields

$$p_a(q) = S'(q) = \ln \left(\frac{q}{2} + \frac{\gamma}{2q} \right). \quad (106)$$

Therefore,

$$S''(q) = \frac{q^2 - \gamma}{q^2(q + \gamma)}, \quad (107)$$

$S''(q_1) = -\delta/(1 - \delta)$, and $S''(q_2) = \delta/(1 + \delta)$. Furthermore, using Eq. (104) we obtain

$$\int_{q_2}^{q_1} \ln \frac{w_-}{w_+} dq = 2 \left(\delta - \sqrt{1 - \delta^2} \arctan \frac{\delta}{\sqrt{1 - \delta^2}} \right) \quad (108)$$

and

$$\exp \left[- \int_{q_2}^{q_1} \left(\frac{u_+}{w_+} - \frac{u_-}{w_-} \right) dq \right] = \sqrt{\frac{1 + \delta}{1 - \delta}}.$$

Plugging everything into Eq. (95) yields the MTE in physical units:

$$\tau = \frac{\pi(1 - \delta)}{\mu\delta} e^{N\Delta S}. \quad (109)$$

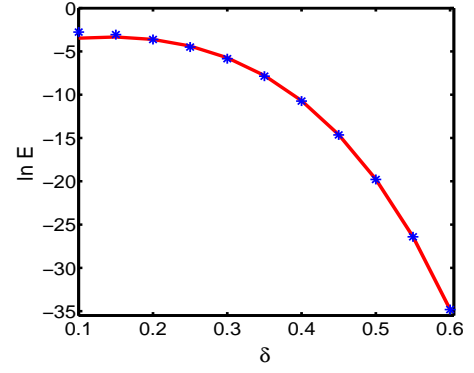


FIG. 6: (color online). The natural logarithm of the extinction rate $E = 1/\tau$ versus δ for $N = 200$ for the reactions $A \xrightarrow{\mu} \emptyset$, $2A \xrightarrow{\lambda} 3A$ and $3A \xrightarrow{\sigma} 2A$. The analytical result (109) (the solid line) is compared with the quantity $-\ln(1 - P_0^{num}(t))/t$ (the asterisks) extracted from a numerical solution of (a truncated) master equation Eq. (2) with rates (103).

Here

$$\Delta S = 2 \left(\delta - \sqrt{1 - \delta^2} \arctan \frac{\delta}{\sqrt{1 - \delta^2}} \right) \quad (110)$$

is a monotone increasing function of δ ; its asymptotes are

$$\Delta S = \begin{cases} (2/3)\delta^3 + (4/15)\delta^5 + \dots, & \delta \ll 1, \\ 2 - \pi\sqrt{2(1 - \delta)} + \dots, & 1 - \delta \ll 1. \end{cases}$$

Near the bifurcation, $\delta \ll 1$, Eq. (109) becomes

$$\tau = \frac{\pi}{\alpha\delta} \exp \left(\frac{2}{3} N \delta^3 \right). \quad (111)$$

This is in agreement with Eq. (102). Indeed, at the bifurcation one has $H_{pp}(1, 0) = 4$, $H_{qp}(1, 0) = -2$, and so $D_B^2 = 1/2$. Equation (111) is valid when $N\delta^3 \gg 1$. This inequality, combined with $\delta \ll 1$, yields the double inequality $N^{-1/3} \ll \delta \ll 1$.

We compared our analytical result (109) with numerical solutions of (a truncated) master equation (2) with rates (103) at $N = 200$ and different values of δ . The comparison is presented in Fig. 6. Very good agreement is observed for not too small δ , when the effective entropy barrier $N\Delta S$ is sufficiently high.

To determine the QSD in this example, one needs to find $S^{(f)}(q)$ and $S_1^{(f)}(q)$ from Eqs. (14) and (17):

$$\begin{aligned} S^{(f)}(q) &= 2\sqrt{\gamma} \arctan \left(\frac{q}{\sqrt{\gamma}} \right) + q \left[\ln \left(\frac{\gamma + q^2}{2q} \right) - 1 \right] \\ S_1^{(f)}(q) &= -\ln \left(\frac{\gamma + q^2}{q^{5/2}} \right). \end{aligned} \quad (112)$$

Then Eqs. (24), (26), (37) and (84) yield the QSD in terms of four overlapping asymptotes. The QSD and its comparison with numerics are shown in Fig. 7, and very good agreement is observed.

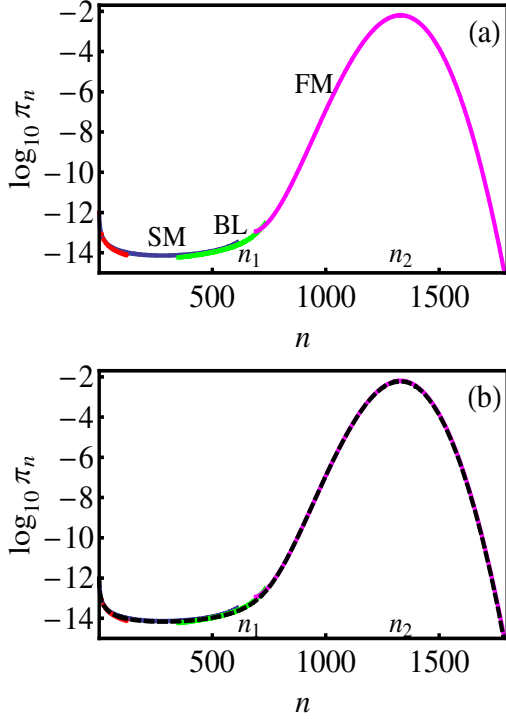


FIG. 7: (color online). (a): the QSD π_n for $\delta = 1/3$ and $N = 10^3$ for the reactions $A \xrightarrow{\mu} \emptyset$, $2A \xrightarrow{\lambda} 3A$ and $3A \xrightarrow{\sigma} 2A$. The QSD includes four overlapping asymptotes: the fast-mode solution (FM), the slow-mode solution (SM), the boundary-layer solution (BL), and the small- n asymptote. (b) a comparison between the QSD in (a) (the solid line) and a numerical solution of (a truncated) master equation (2) with rates (103) (the dashed line).

VI. SUMMARY

This work dealt with extinction of an isolated long-lived stochastic population describable by a continuous-time Markov process. We have identified two generic extinction scenarios, A and B, based on the stability properties of the fixed points of the population size dynamics, predicted by the rate equation. For each of the two scenarios we have calculated the mean time to extinction (MTE) and the quasi-stationary probability distribution (QSD). The calculations involve a systematic use of WKB method, where $1/N$, the typical inverse population size in the metastable state, serves as a small parameter of the theory. The WKB theory is supplemented by two additional approximations in the regions where it breaks down. One of them is a small- n expansion of the quasi-stationary master equation which brings about a recursion relation (in both scenarios A and B). The second one is the Fokker-Planck equation, obtained via the van Kampen system-size expansion. The latter is valid in small regions around the non-trivial fixed points of the rate equation (in scenario B).

The theory is not limited to single-step stochastic pro-

cesses, although for such processes our general results simplify considerably. The results also simplify near the characteristic bifurcations of scenarios A and B. A number of previous results for the mean time to extinction follow from our equations in particular cases.

We have observed that, in models belonging to extinction scenario B, the mean time to population extinction formally coincides, in the leading and subleading orders in N , with the mean time to population *escape* in stochastic population models where the metastable population switches to another metastable state or to an absorbing state at infinity. The reason for this coincidence is that, in all these systems, a constant probability current sets in beyond the repelling fixed point of the rate equation. It is the magnitude of this current, rather than the precise nature of the target state for escape, which determines, in the leading and subleading orders in N , the mean escape rate from a metastable state.

The situation is quite different in extinction scenario A. Here the exact nature of the target state (the absorbing state at $n = 0$) or, more precisely, the rate constants of the effective linear branching processes at small n , affects the pre-exponent of the mean time to extinction.

Acknowledgments

We are very grateful to B. Derrida, M. Dykman, A. Kamenev, and P.V. Sasorov for discussions. B. M. and M. A. were supported by the Israel Science Foundation (Grant No. 408/08). M. A. was supported by the Clore Foundation.

Appendix A

Here we derive Eq. (35) for the arbitrary constants C_i appearing in Eq. (34). For that we need to solve the following set of $K + 1$ linear equations:

$$C_0 + \sum_{i=1}^K \lambda_i^{-m} C_i = f_m, \quad m = 1, 2, \dots, K + 1, \quad (\text{A1})$$

where the constants f_2, f_3, \dots, f_{K+1} can be expressed via f_1 by using the recursive Eq. (30) with $n = 1, 2, \dots, K$ (assuming that $f_j = 0$ for $j < 1$).

As before, we denote the roots of Eq. (31) by $\lambda_0, \lambda_1, \dots, \lambda_K$, where $\lambda_0 = 1$, and define the following quantities:

$$\begin{aligned} s_0^{(0)} &= 1 \\ s_1^{(0)} &= \lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_K \\ s_2^{(0)} &= \lambda_1 \lambda_2 + \lambda_2 \lambda_3 + \lambda_3 \lambda_4 + \dots + \lambda_{K-1} \lambda_K \\ s_3^{(0)} &= \lambda_1 \lambda_2 \lambda_3 + \lambda_1 \lambda_3 \lambda_4 + \lambda_1 \lambda_2 \lambda_4 + \dots \\ &\dots\dots\dots \\ s_K^{(0)} &= \lambda_1 \lambda_2 \dots \lambda_K. \end{aligned} \quad (\text{A2})$$

The superscript (0) means that the K roots which contribute to $s_i^{(0)}$ are all of the roots of Eq. (31) *except* λ_0 . In the same manner we can define for $1 \leq i \leq K$

$$\begin{aligned} s_0^{(i)} &= 1 \\ s_1^{(i)} &= 1 + \lambda_1 + \lambda_2 + \cdots + \lambda_{i-1} + \lambda_{i+1} + \cdots + \lambda_K \\ s_2^{(i)} &= \sum_{\substack{0 \leq j < m \leq K \\ j, m \neq i}} \lambda_j \lambda_m \\ &\dots\dots \\ s_K^{(i)} &= \lambda_1 \lambda_2 \cdots \lambda_{i-1} \lambda_{i+1} \cdots \lambda_K, \end{aligned} \quad (\text{A3})$$

where the root λ_i does *not* contribute.

To obtain the coefficient C_0 we multiply the first of Eqs. (A1) by $s_0^{(0)}$, the second by $-s_1^{(0)}$, ..., and finally the last by $(-1)^K s_K^{(0)}$. By adding the equations we obtain

$$\begin{aligned} &f_1 - s_1^{(0)} f_2 + s_2^{(0)} f_3 - \cdots + (-1)^K s_K^{(0)} f_{K+1} \\ &= C_0 \left[1 - s_1^{(0)} + s_2^{(0)} - \cdots + (-1)^K s_K^{(0)} \right] \\ &+ C_1 \left[\lambda_1^{-1} - \lambda_1^{-2} s_1^{(0)} + \lambda_1^{-3} s_2^{(0)} - \cdots + (-1)^K \lambda_1^{-K-1} s_K^{(0)} \right] \\ &+ \cdots + C_K \left[\lambda_K^{-1} - \lambda_K^{-2} s_1^{(0)} + \cdots + (-1)^K \lambda_K^{-K-1} s_K^{(0)} \right]. \end{aligned} \quad (\text{A4})$$

By using Eqs. (A2) we can rewrite the coefficient of C_0 in the right hand side of Eq. (A4) as

$$1 - s_1^{(0)} + s_2^{(0)} - \cdots + (-1)^K s_K^{(0)} = (1 - \lambda_1)(1 - \lambda_2) \cdots (1 - \lambda_K). \quad (\text{A5})$$

Furthermore, the coefficient of C_i in Eq. (A4) satisfies

$$\begin{aligned} &\lambda_i^{-1} - \lambda_i^{-2} s_1^{(0)} + \lambda_i^{-3} s_2^{(0)} - \cdots + (-1)^K \lambda_i^{-K-1} s_K^{(0)} \\ &= \frac{(\lambda_i - \lambda_1)(\lambda_i - \lambda_2) \cdots (\lambda_i - \lambda_j) \cdots (\lambda_i - \lambda_K)}{\lambda_i^{K+1}}. \end{aligned} \quad (\text{A6})$$

Clearly, this expression vanishes for all $i \geq 1$. Therefore, Eqs. (A4) and (A5) yield

$$C_0 = \frac{f_1 - s_1^{(0)} f_2 + s_2^{(0)} f_3 - \cdots + (-1)^K s_K^{(0)} f_{K+1}}{\prod_{i=1}^K (1 - \lambda_i)} \quad (\text{A7})$$

To calculate the numerator of Eq. (A7) we use Viète's formula. Given a polynomial equation

$$a_K x^K + \cdots + a_1 x + a_0 = 0,$$

whose roots are $\lambda_1, \lambda_2, \dots, \lambda_K$, the expressions $s_i^{(0)}$, given by Eq. (A2), satisfy

$$s_i^{(0)} = (-1)^i \frac{a_{K-i}}{a_K}. \quad (\text{A8})$$

Let us apply this formula to Eq. (33) which has exactly the roots $\lambda_1, \lambda_2, \dots, \lambda_K$. For convenience, we rewrite Eq. (33) as

$$1 - \lambda \sum_{r=1}^K w'_r(0) - \lambda^2 \sum_{r=2}^K w'_r(0) - \cdots - \lambda^K w'_K(0) = 0. \quad (\text{A9})$$

In terms of the coefficients a_i we have $a_0 = 1$, $a_1 = -[w'_1(0) + \cdots + w'_K(0)]$, $a_2 = -[w'_2(0) + \cdots + w'_K(0)]$, ..., and $a_K = -w'_K(0)$. Therefore, using Eqs. (A7) and (A8), we can rewrite C_0 as

$$C_0 = \frac{f_{K+1} - f_K[w'_1(0) + \cdots + w'_K(0)] - \cdots - f_1 w'_K(0)}{a_K \prod_{i=1}^K (1 - \lambda_i)}. \quad (\text{A10})$$

Now, using Eq. (30) with $n = K$, we obtain a relation between f_{K+1} and $f_{j \leq K}$. Plugging it into (A10) we have

$$\begin{aligned} C_0 &= \frac{1}{a_K \prod_{i=1}^K (1 - \lambda_i)} \{f_K - f_{K-1}[w'_1(0) + \cdots + w'_K(0)] \\ &- \cdots - f_1[w'_{K-1}(0) + w'_K(0)]\}. \end{aligned} \quad (\text{A11})$$

One can use this argument repeatedly $K - 1$ more times, and obtain that the expression in the curly brackets in Eq. (A11) equals f_1 . In addition, by virtue of Eq. (A8) $w'_K(0)$ satisfies

$$w'_K(0) = -a_K = \frac{(-1)^{K+1}}{\lambda_1 \lambda_2 \cdots \lambda_K}. \quad (\text{A12})$$

Therefore, C_0 is given by

$$C_0 = \frac{(-1)^K f_1 \prod_{i=1}^K \lambda_i}{\prod_{i=1}^K (1 - \lambda_i)} = \frac{f_1}{\prod_{i=1}^K (1 - \lambda_i^{-1})}, \quad (\text{A13})$$

thereby proving Eq. (35) for $i = 0$.

This proof can be generalized to the rest of the coefficients C_i . Here one has to multiply the first of Eqs. (A1) by $s_0^{(i)}$, the second by $-s_1^{(i)}$, ..., and finally the last by $(-1)^K s_K^{(i)}$, and add all the equations. This yields

$$\begin{aligned} &f_1 - s_1^{(i)} f_2 + s_2^{(i)} f_3 - \cdots + (-1)^K s_K^{(i)} f_{K+1} \\ &= C_0 \left[1 - s_1^{(i)} + s_2^{(i)} - \cdots + (-1)^K s_K^{(i)} \right] \\ &+ C_1 \left[\lambda_1^{-1} - \lambda_1^{-2} s_1^{(i)} + \lambda_1^{-3} s_2^{(i)} - \cdots + (-1)^K \lambda_1^{-K-1} s_K^{(i)} \right] \\ &+ \cdots + C_K \left[\lambda_K^{-1} - \lambda_K^{-2} s_1^{(i)} + \cdots + (-1)^K \lambda_K^{-K-1} s_K^{(i)} \right]. \end{aligned} \quad (\text{A14})$$

The coefficient of C_i in the right hand side of Eq. (A14) satisfies

$$\begin{aligned} & \lambda_i^{-1} - \lambda_i^{-2} s_1^{(i)} + \lambda_i^{-3} s_2^{(i)} - \cdots + (-1)^K \lambda_i^{-K-1} s_K^{(i)} \\ &= \frac{(\lambda_i - 1)(\lambda_i - \lambda_1) \cdots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \cdots (\lambda_i - \lambda_K)}{\lambda_i^{K+1}}, \end{aligned} \quad (\text{A15})$$

because λ_i is absent from $s_j^{(i)}$, see Eq. (A3). On the other hand, the coefficients of all other $C_{j \neq i}$ in the right hand side of Eq. (A14) can be shown to be equal zero. Therefore,

$$C_i = \frac{\lambda_i^{K+1} \left[f_1 - s_1^{(i)} f_2 + s_2^{(i)} f_3 - \cdots + (-1)^K s_K^{(i)} f_{K+1} \right]}{\prod_{\substack{j=0 \\ j \neq i}}^K (\lambda_i - \lambda_j)} \quad (\text{A16})$$

By using the recursive equation (30) repeatedly and by using Eq. (A3), one finally obtains Eq. (35).

Finally, C_0 from Eq. (A13) can be expressed through the reaction rate constants $w'_1(0), w'_2(0), \dots, w'_K(0)$, see Eq. (36). Indeed, let us expand the denominator of Eq. (A13) and use Eq. (A2):

$$\begin{aligned} & (1 - \lambda_1^{-1})(1 - \lambda_2^{-1}) \cdots (1 - \lambda_K^{-1}) = 1 - \sum_i \lambda_i^{-1} \\ & + \sum_{i>j} \lambda_i^{-1} \lambda_j^{-1} + \cdots + (-1)^K \lambda_1^{-1} \cdots \lambda_K^{-1} \\ &= \frac{(-1)^K}{\lambda_1 \cdots \lambda_K} \left[(-1)^K s_K^{(0)} + (-1)^{K-1} s_{K-1}^{(0)} + \cdots + 1 \right]. \end{aligned} \quad (\text{A17})$$

Using Eqs. (A8) and (A12), we rewrite Eq. (A17) as

$$\begin{aligned} & (1 - \lambda_1^{-1})(1 - \lambda_2^{-1}) \cdots (1 - \lambda_K^{-1}) = 1 - [w'_1(0) + \cdots \\ & + w'_K(0)] - [w'_2(0) + \cdots + w'_K(0)] - \cdots - w'_K(0) \\ &= 1 - w'_1(0) - 2w'_2(0) - 3w'_3(0) - \cdots - Kw'_K(0). \end{aligned} \quad (\text{A18})$$

Plugging this into Eq. (A13) one obtains Eq. (36).

Appendix B

Here we show that, in extinction scenario A, the two real and positive roots of Eq. (31) are $\lambda_0 = 1$ and $0 < \lambda_1 < 1$, whereas all other roots obey the inequality $|\lambda_{i>1}| > \lambda_1$. We start by showing that the positive root of Eq. (33) obeys the inequalities $0 < \lambda_1 < 1$. The left hand side of Eq. (33) side is a monotone decreasing function of λ . At $\lambda = 1$ it is equal to $1 - w'_1(0) - 2w'_2(0) - \cdots - Kw'_K(0)$. This quantity is negative, as $n = 0$ is a repelling fixed point, so the rescaled reaction rate constants satisfy the inequality

$w'_1(0) + 2w'_2(0) + \cdots + Kw'_K(0) - 1 > 0$. On the other hand, at $\lambda = 0$ the left hand side is 1. Hence, $0 < \lambda_1 < 1$.

Now we will prove by contradiction that all other (negative or complex) roots satisfy the inequality $|\lambda_{i>1}| > \lambda_1$. Assume by contradiction that there exists a root λ_j so that $|\lambda_j| \leq \lambda_1$. Denote $\lambda_j = ae^{i\theta}$. Then by assumption $a \leq \lambda_1$. Substituting λ_j into Eq. (33) we have

$$\begin{aligned} & 1 - w'_1(0)\lambda_j - \cdots - w'_K(0)(\lambda_j + \cdots + \lambda_j^K) \\ &= 1 - w'_1(0)a \cos \theta - \cdots \\ & - w'_K(0)(a \cos \theta + \cdots + a^K \cos K\theta) + i(\cdots) = 0, \end{aligned} \quad (\text{B1})$$

where both real and imaginary parts have to vanish separately. Now we substitute λ_1 into Eq. (33) and use Eq. (B1)

$$\begin{aligned} 0 &= 1 - w'_1(0)\lambda_1 - \cdots - w'_K(0)(\lambda_1 + \cdots + \lambda_1^K) \\ &\leq 1 - w'_1(0)a - \cdots - w'_K(0)(a + \cdots + a^K) \\ &< 1 - w'_1(0)a \cos \theta - \cdots \\ & - w'_K(0)(a \cos \theta + \cdots + a^K \cos K\theta) = 0, \end{aligned} \quad (\text{B2})$$

where the last inequality holds since λ_j is complex or negative, so $\theta \neq 0$, and there exists some m for which $\cos m\theta < 1$.

Eq. (B2) shows a contradiction $0 < 0$, so all roots obey $|\lambda_{i>1}| > \lambda_1$. As a result, at $n \gg 1$, the recursive solution (37) reduces to Eq. (38), where A_1 is given by Eq. (39).

Finally, we show that $A_1 > 0$, and so the asymptote (38) is always positive. First, by using Eq. (A12), one can see that the numerator in Eq. (39) is always negative. What is the sign of the denominator? Here $\lambda_1 - \lambda_0 < 0$, whereas all other terms in the product are positive. Indeed, for $\lambda_j < 0$ one has $\lambda_1 - \lambda_j > 0$. For any complex λ_j , there is also a complex conjugate root $\lambda_k = \bar{\lambda}_j$. Therefore, by writing $\lambda_j = a + ib$, one has $(\lambda_1 - \lambda_j)(\lambda_1 - \bar{\lambda}_j) = (\lambda_1 - a)^2 + b^2 > 0$. So, A_1 is always positive, and so is the asymptote $\pi(q)$ from Eq. (38).

Appendix C

Here we show that, in extinction scenario B, the root $\lambda_0 = 1$ of Eq. (31) has the smallest absolute value among all of the roots. To this end we will prove that the roots of Eq. (33) obey the inequality $|\lambda_{j>0}| > 1$. Let us denote $\lambda_{j>0} = a_j e^{i\theta_j}$, in general a complex number, and assume by contradiction that $a_j \leq 1$. Plugging $\lambda = \lambda_j$ into Eq. (33), we obtain

$$\begin{aligned} & 1 - a_j \cos(\theta_j)w'_1(0) - [a_j \cos(\theta_j) + a_j^2 \cos(2\theta_j)]w'_2(0) - \cdots \\ & - [a_j \cos(\theta_j) + \cdots + a_j^K \cos(K\theta_j)]w'_K(0) + i(\cdots) = 0, \end{aligned} \quad (\text{C1})$$

where the real and imaginary parts must vanish separately. As we have assumed $a_j \leq 1$, we have $a_j^m \cos(m\theta_j) \leq 1$ for all integer $m \geq 0$. Therefore, we

can write for the real part of Eq. (C1):

$$\begin{aligned} 0 &= 1 - a_j \cos(\theta_j) w'_1(0) - [a_j \cos(\theta_j) + a_j^2 \cos(2\theta_j)] w'_2(0) \\ &\quad - \cdots - [a_j \cos(\theta_j) + \cdots + a_j^K \cos(K\theta_j)] w'_K(0) \\ &\geq 1 - w'_1(0) - 2w'_2(0) - \cdots - K w'_K(0) > 0, \end{aligned} \quad (\text{C2})$$

where the last inequality follows from $n = 0$ being an attracting fixed point of the rate equation. Equation (C2) shows a contradiction $0 > 0$. Hence $a_j > 1$, and all the roots of Eq. (33) obey the inequality $|\lambda_{j>0}| > 1$.

-
- [1] M.S. Bartlett, *Stochastic Population Models in Ecology and Epidemiology* (Wiley, New York, 1961).
 - [2] S. R. Beissinger and D. R. McCullough (Editors), *Population Viability Analysis* (University of Chicago Press, Chicago, 2002).
 - [3] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*, Lect. Notes Stat., Vol. **151** (Springer, New York, 2000).
 - [4] M.S. Samoilov and A.P. Arkin, *Nature Biotech.* **24**, 1235 (2006); M. Assaf and B. Meerson, *Phys. Rev. Lett.* **100**, 058105 (2008).
 - [5] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 2001).
 - [6] C.W. Gardiner, *Handbook of Stochastic Methods* (Springer Verlag, Berlin, 2004).
 - [7] We will assume that the stochastic population does not exhibit an unlimited growth (escape to infinite population size), see Ref. [21], and $n = 0$ is the *only* absorbing state.
 - [8] M. Assaf and B. Meerson, *Phys. Rev. E* **74**, 041115 (2006).
 - [9] M. Assaf and B. Meerson, *Phys. Rev. Lett.* **97**, 200602 (2006).
 - [10] C.M. Bender and S.A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (Springer, New York, 1999).
 - [11] R. Kubo, K. Matsuo, and K. Kitahara, *J. Stat. Phys.* **9**, 51 (1973).
 - [12] M.I. Dykman, E. Mori, J. Ross, and P.M. Hunt, *J. Chem. Phys.* **100**, 5735 (1994).
 - [13] V. Elgart and A. Kamenev, *Phys. Rev. E* **70**, 041106 (2004).
 - [14] V. Elgart and A. Kamenev, *Phys. Rev. E* **74**, 041101 (2006).
 - [15] M. Assaf, A. Kamenev, and B. Meerson, *Phys. Rev. E* **78**, 041123 (2008).
 - [16] D.A. Kessler and N.M. Shnerb, *J. Stat. Phys.* **127**, 861 (2007).
 - [17] B. Gaveau, M. Moreau, and J. Toth, *Lett. Math. Phys.* **37**, 285 (1996).
 - [18] C.R. Doering, K.V. Sargsyan, and L.M. Sander, *Multi-scale Model. and Simul.* **3**, 283 (2005).
 - [19] M. Assaf and B. Meerson, *Phys. Rev. E* **75**, 031122 (2007).
 - [20] J.W. Turner and M. Malek-Mansour, *Physica A* **93**, 517 (1978).
 - [21] B. Meerson and P.V. Sasorov, *Phys. Rev. E* **78**, 060103(R) (2008).
 - [22] C. Escudero and A. Kamenev, *Phys. Rev. E* **79**, 041149 (2009).
 - [23] At nontrivial fixed points of the rate equation, $q = q_i = n_i/N > 0$, the two real roots of the equation $H(q, p) = 0$ merge at $p = 0$.
 - [24] If $S''(0) = 0$, then it is more convenient to use Eq. (17).
 - [25] J.N. Darroch and E. Seneta, *J. Appl. Probab.* **4**, 192 (1967).
 - [26] We remind the reader that, in order to obtain Eq. (9), we rescaled the reaction rates and time by the linear decay rate constant $w'_{-1}(0) = \alpha$. To express the MTE in physical units one needs to put the factor α back.
 - [27] M. Assaf, A. Kamenev, and B. Meerson, *Phys. Rev. E* **79**, 011127 (2009).
 - [28] I. Näsell, *J. Theor. Biol.* **211**, 11 (2001).
 - [29] O. Ovaskainen, *J. Appl. Prob.* **38**, 898 (2001).
 - [30] In Refs. [28, 29] the parameter R_0 was defined as $\lambda(N - 1)/\mu$. Therefore, their result for the MTE looks slightly different, but it is actually identical to Eq. (71) in the leading and subleading orders of N .
 - [31] M.I. Dykman and M.A. Krivoglaz, *Physica A* **104**, 480 (1980).
 - [32] The distance to bifurcation δ_1 is defined in Refs. [12] and [31] differently than in the present work. Their parameter δ_1 is related to our parameter δ as $\delta_1 = \delta^2$, so the entropy barriers in Refs. [12] and [31] scale with δ_1 as $\delta_1^{3/2}$.